NASA-CR-190894

SBIR - 06.05-4100
release date - 11/09/88

Technical Report
1977

OCULAR ATTENTION-SENSING INTERFACE SYSTEM

December 1986

Allen Zaklad, Ph.D.
Floyd A. Glenn III, Ph.D.
Helene P. Iavecchia
James M. Stokes

Submitted to:

NASA Resident Office – JPL
Attn: Chief, Procurement Branch
4800 Oak Grove Drive
Pasadena, CA 91109

Final Report
Contract No. NAS7-932

(NASA-CR-190884) OCULAR ATTENTION-SENSING INTERFACE SYSTEM Final Technical Report, 1977 (Analytics) 76 p

N93-13450

Unclas

G3/54 0121203

ANALYTICS

2500 MARYLAND RD., WILLOW GROVE, PA. 19090

PROJECT SUMMARY: OCULAR ATTENTION-SENSING INTERFACE SYSTEM (OASIS)
PHASE II SBIR, CONTRACT NAS7-932

The purpose of the research was to develop an innovative human-computer interface based on eye movement and voice control. By eliminating a manual interface (keyboard, joystick, etc.), OASIS provides a control mechanism that is natural, efficient, accurate, and low in workload.

The research covered four overall tasks:

1. Building a laboratory facility for OASIS interface experimentation and system development.
2. Determining the characteristics of an optimal eye-voice interface. This step comprises both informal investigation and full-scale controlled human-system performance studies.
3. Demonstrating the utility of this interface for typical applications involving tactical display interaction. These simulated tasks include remote manipulation, product inspection, targeting and firing, and multiple vehicle control.
4. Preparing engineering development plans for the final OASIS design.

Tasks 1, 2, and 3 were completed. Using the laboratory testbed, modes of visual feedback and filtering algorithms were studied experimentally. The results showed that the OASIS interface has great future potential in a number of possible applications. In addition, the OASIS interface compared favorably with the conventional rapid-pointing mouse interface for the tasks studied. More extensive experimentation is recommended to optimize the interface. Task 4 was begun and was not completed since the final design of the interface will be driven by the requirements of a specific application.

Potential commercial and governmental applications have been identified, though no commitments have been made. These application areas include tactical targeting and system control, intention sensing, loss of consciousness determination, rapid database manipulation, robotic control, and aids for the physically disabled.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

FIGURES

ACKNOWLEDGMENTS

The work reported here was performed under NASA's Small Business Innovative Research (SBIR) program (Contract Number NAS7-932). Monitoring of the contract was performed by the NASA Resident Office at the Jet Propulsion Laboratory (JPL). We gratefully acknowledge the assistance we received from Mr. Peter Tackney, Contracting Officer, and Mr. Marvin Perlman, Contracting Officer's Technical Representative, at JPL.

Many people at NASA centers offered their assistance to the OASIS project, especially Dr. Randall Harris (Langley), Dr. Mel Montemerlo (HQ), and Dr. Steve Ellis (Ames). Dr. Felix Barker and his associates at the Pennsylvania College of Optometry provided special clinical expertise. We also want to thank the two major equipment vendors -- Applied Science Laboratories (oculometer) and Masscomp (graphics and system integration) -- for timely help in troubleshooting our system.

Finally, we wish to acknowledge the many people at Analytics who helped us reach a substantial degree of success in building OASIS. Bill Weiland, Lorna Ross, Dan Leibholz, and Dan Weiss played important roles in designing and building the software. Steve Rodgers and Ed Kapnic helped us to bring OASIS into the public eye. Steve Leibholz, the President of Analytics, contributed both technical ideas and inspiration to the project.

1. INTRODUCTION AND OVERVIEW

This report documents and describes the research conducted in the SBIR Phase II program, under contract NAS7-932, on the Ocular Attention-Sensing Interface System (OASIS).

1.1 INTRODUCTION

The manner by which human users communicate with computers is widely recognized as a key delimiter of the usefulness of computers and is an underlying consideration in the conception of OASIS. "User friendliness" of user-computer interfaces is generally regarded as a highly desirable but elusive feature, with most computer systems requiring the user/operator to learn complex, unnatural protocols associated with keyboards, programming languages, and graphic interaction devices such as the joystick or trackball. While it may be possible to learn the principles of operation of these interface devices in a tolerable amount of time, it takes a prolonged period of practice to become facile with any of them. In many systems, these devices already exist, and added workload would be generated by the incorporation of yet another manual interface. This is not feasible and would hamper incisive action on the part of the operator. The use of OASIS eliminates the need for additional manual input devices.

The two principal domains of human-computer communication are continuous, spatial information and discrete, verbal information. The most natural means for a human to interact with a spatial scene is to direct attention through eye movements and, with verbal information, analogously, to speak and listen in a natural-language medium. Current technologies permit a human's eye movements and fixations to be automatically tracked using devices known as eye trackers or oculometers. Human speech, with constrained vocabularies and syntax structures, can be automatically recognized and interpreted by a variety of commercially available devices, and very effective speech generation devices are also widely available. However, a useful user-computer interface cannot be constructed simply by connecting an eye-tracking device and an automated speech processor to a computer. The objective can only be met by implementation of a special process to deduce human cognition and decision. This is OASIS.

1.1.1 The OASIS Interface

The OASIS interface is an innovative concept, composed principally of an automated eye-tracking system and a speech recognizer, that directs computer resources based on an observer's visual attention. Computer systems that recognize human intentions and human language provide a powerful communication channel never before realized.

The human eye is an excellent pointer. A person can look at objects of interest directly and steadily. The eye can also examine fine details in the visual field, look away, and return to any fine detail with swiftness and accuracy. This can be done again and again very reliably. Due to the nature of eye movement, it is less clear how well the eye can perform a control task. The eyes can function as a quasi-guidance system; they provide valuable assistance when a person is moving through space. The eyes, however, do not move smoothly; they move in a jerky manner and are constantly in motion. When a person looks at an object in the visual field, he is not aware of his own eye movement activity because he sees stable images. This stability is really the result of many micromovements of the eyes.

To control a cursor on a display screen, the eyes need a supplementary system for fine tuning. For example, if the user of the system is visually tracking an object on the display screen and would like the display screen cursor to be on the targeted object, a voice command is used to center the cursor on the targeted object. The system user employs natural-language commands that fine tune the cursor control. Speech recognizers employ a pattern-matching technique to compare a spoken word to a collection of stored-word reference patterns. Concurrent information, such as eye position, can amplify and clarify what is said. Together, eye movements and voice commands are combined in OASIS to control complex systems.

The first element in the development of OASIS is the determination of the focus of human visual attention from a stream of measurements of eye position and orientation, eliminating noise and bias and introducing temporal adjustment. The second element is the incorporation, via conventional voice input processors, of voice direction to advise the attention-estimating system of the significance of human attention. Voice input data are correlated with visual-processed data via a time lag adjustment which takes into account saccadic eye movement. These are necessary because human attention may be directed at several objects within view at the same time, and the command delay is a function of cognitive workload. The interaction of voice command to the system with attention-focus processing yields results not obtainable with a simple combination of such inputs.

The operator's attention position, as opposed to the eye position, is the primary source of data for OASIS. The attention position is a more accurate indicator of the operator's visual point of interest due to elements which cause variations in the exact location of the eye position. These elements include:

- the tracking of multiple objects,
- blinking,
- major saccades (e.g., caused by eye movements),
- minor saccades (e.g., small jumps, settling time, noise in the muscle),
- small tracking deviations due to the eye position lagging behind head movements, and
- damped muscular oscillation after a saccade.

The attention position is established through filtering algorithms and the use of voice input data. A digital filter is used to reject unwanted signals and select the eye motion data related to visual foci.

1.2 BACKGROUND -- PHASE I

The OASIS project represents the completion of SBIR Phases I and II (contracts NAS7-922 and NAS7-932, respectively). Before presenting the Phase II research, a brief summary of the Phase I effort (Glenn et al., 1984) is given.

The overall objective of the Phase I effort was to assess the technical feasibility and viability of the OASIS interface. This assessment was carried out via the following specific technical tasks:

- develop functional specifications for the OASIS concept;
- determine the technology status of current voice and eye-tracking technology;
- address the human factors issues of an OASIS interface, including feedback, coordination, timing, fatigue, and stress;
- identify the generic characteristics of applications suitable for OASIS;
- recommend specific application areas; and
- develop an experimental plan to (a) conduct basic research for optimizing the OASIS interface, and (b) demonstrate the utility of OASIS by simulating various application areas.

All of these technical objectives for the Phase I effort were successfully met. Our findings have allowed us to explore the technology requirements of OASIS, and we have determined ways to integrate the technology and construct a prototype system. As a result, in our Phase I final report, we were able to show the technical feasibility of OASIS.

1.2.1 OASIS Functional Architecture

A functional overview for a prototype OASIS system for experimental use is shown in Figure 1-1. Input into that system can originate from conventional

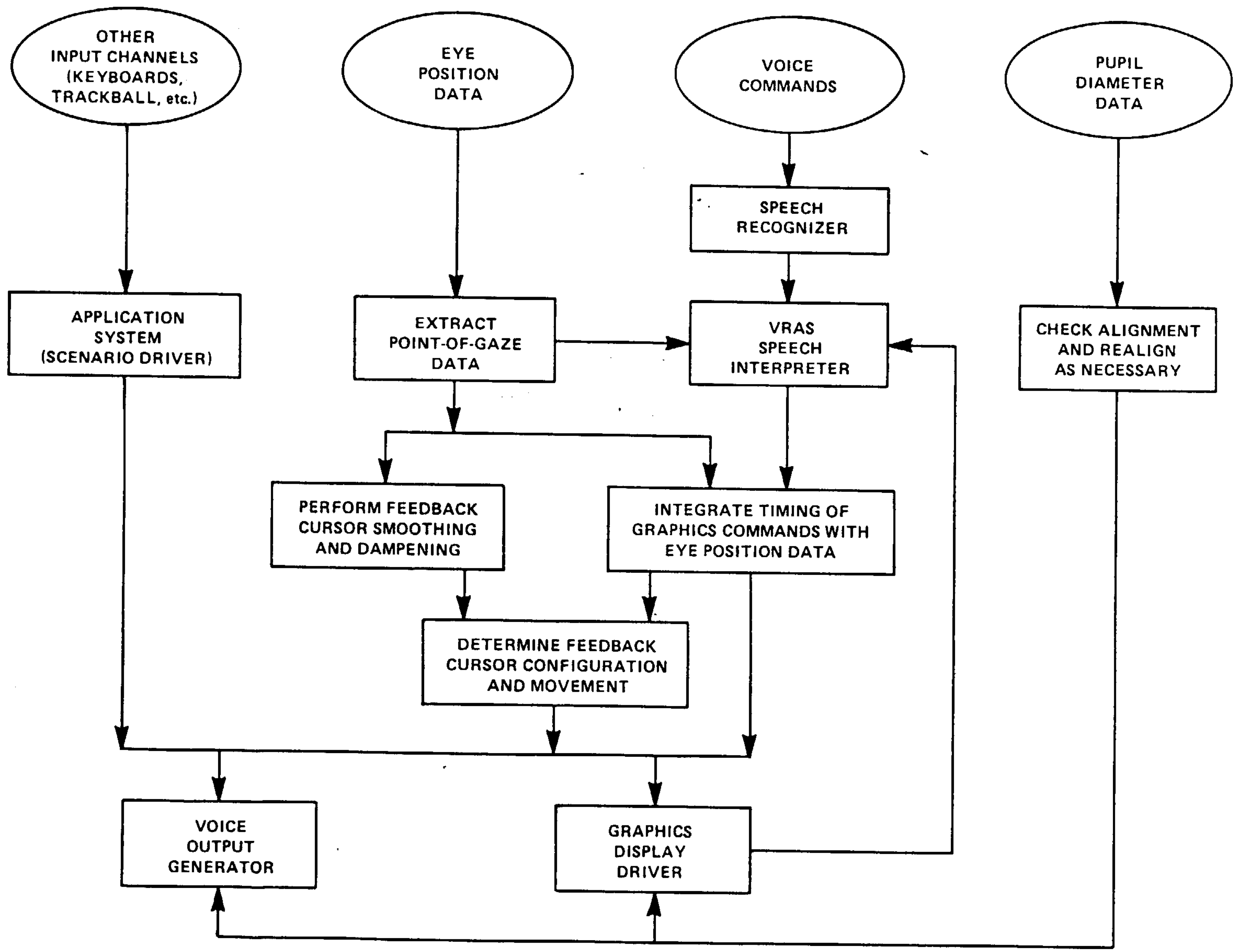


Figure 1-1. Functional Overview for Prototype OASIS System

86-21-07-FR

input devices operated by the subject or the experimenter (such as a keyboard or mouse), from the oculometer, and from the voice recognizer. Subject-initiated keyboard or mouse commands may be used in combination with OASIS inputs or in circumstances in which the subject's performance, using only conventional input devices, is being recorded. Experimenter-initiated keyboard or mouse commands may be utilized to initialize trial parameters or to dynamically modify algorithm parameters during trial runs.

Rapid and precise timing control is required for the OASIS experimental prototype. The oculometer will send the subject's eye position and pupil diameter to the system 60 times a second. It is unnecessary for the feedback cursor to maintain a similar update rate. Oculometer output points can be averaged and dampened to create the most appropriate dynamics for effective cursor control. The pupil diameter data can serve as a coarse indicator of operator workload for both conventional and OASIS interface configurations. Pupil diameter can also serve as a flag to indicate loss of the pupil image on the oculometer camera due to blinks or head movement when the diameter is continuously zero.

1.2.2 Eye Tracking and Voice Technology Status

1.2.2.1 Eye-Tracking Technology Status. There are several methods of tracking eye movements, but the results of the Phase I research effort showed conclusively that only the corneal reflection method is appropriate for OASIS application. The corneal reflection technique uses remote equipment which does not require attachment to the subject. Some corneal reflection-based systems allow some free head movement, thus precluding the need for a bite board or chin rest which would interfere with production of voice commands.

In general, the Phase I research effort recommends an eye-tracking system for OASIS that meets the following criteria:

- It must be non-intrusive (no contact lenses, electrodes, etc.).
- It must be non-restrictive (speech must not be impeded and some free head movement allowed).
- It must be accurate to within at least one degree of visual angle within a field-of-view ranging over at least 20 degrees both vertically and horizontally.
- It must be fully automatic and provide eye position data in real time.

An investigation of manufacturers revealed that there is only one commercial source for an oculometer that meets the above criteria; this source is Applied Science Laboratories (ASL). The ASL system employs a TV camera to provide a close-up of the subject's eye with an option to use a computer-controlled moving mirror to hold the image of the eye as the subject's head moves. The TV image of the eye is automatically processed to determine the location of the pupil and hence of the pupil center. A fine unobtrusive infrared light beam is also projected onto the subject's eye, and the reflection of the cornea of the eye is received by the TV camera as a spot superimposed on the pupil. By performing image processing of the relative positions of the pupil and the corneal reflection, it is possible to determine the orientation of the eye and hence where the subject's gaze is directed. Various ASL systems permit different amounts of latitude in the subject's head movement -- up to a maximum allowance of about one cubic foot -- and with all variations providing an accuracy of about one degree of visual angle throughout a large visual field. Although the cost of this type of equipment is currently rather high (a complete, full-featured system with maximum free head movement costs in excess of $150,000), it is expected that substantially reduced costs would result from expanded applications of oculometers as envisioned with OASIS.

1.2.2.2 Voice Technology Status. Automated voice technology has made significant progress recently and, as a result, it now presents a viable control input device characterized by speaker independent recognition of continuous speech, decreased sensitivity to background noise, and increasing levels of recognition accuracy. The recognizer market has been dominated by discrete speech recognizers which process words (or connected phrases treated as words) with a maximum utterance duration between 1 and 1.5 seconds and which must be separated by a period of silence of about 200 milliseconds. The continuous speech recognizers are not limited by utterance length and do not require a pause between words. The OASIS system, then, will utilize a continuous speech recognizer. The vocabulary will effectively be handled through the use of customized software modules which sequence through a hierarchical set of voice commands. In general, discrete and continuous recognition devices can handle vocabularies of a maximum of 100 words (or phrases treated as units) and are capable of greater than 95 percent recognition accuracy on specific vocabularies in benign environments.

1.2.3 Human Factors Issues in OASIS

The primary impetus for the OASIS concept is the naturalness of the voice and eye movement channels for interacting with symbolic graphic information; human thought translates quite naturally and effortlessly into speech and eye movements. At the same time, speech and eye movements are complex control channels with many characteristics that define performance capabilities pertinent to the OASIS concept. Very few of these characteristics can easily be quantitatively specified, however, because of their various complex interactions with task and equipment factors. Categories of performance that deserve particular consideration are voice control, eye control, coordination, stress, and fatigue.

Although speech is a natural medium of expression for humans, speaking so as to be understood by a machine is not. Human listeners are extremely forgiving in accommodating variations and ambiguities in spoken language which cause great difficulty for automatic speech recognition equipment. More than anything else, it is the unpredictable variability in human speech patterns that limits the accuracy of speech recognition technology. This problem exists because we do not know, in a precise physical sense, how to characterize the classes of acoustic patterns that we associate with the words in our languages.

There are large differences in both voice recognition systems and human speakers in dealing with this problem. Some recognition equipment allows much greater latitude than other equipment in enunciation variations. And while there are speakers who are extremely consistent in their enunciations and have little difficulty with any recognition devices, others (the so-called "goats") are extremely variable in their speech and are not accurately recognized by any current devices.

Factors such as stress and fatigue produce changes in individual voices that further complicate the situation. One particularly problematic situation that arises with stress is that stress tends to alter the voice so that misrecognitions occur which, in time-pressured tasks, tend to increase the stress level and hence the likelihood of further misrecognitions. In addition to the physical aspects of voice production, the cognitive aspects of speech also figure importantly in the performance of automated voice systems. Just as voice recognition devices place severe constraints on acoustic characteristics of speech, they also impose constraints on human memory to manage vocabulary and syntax. In this area, a trade-off must be recognized between complexity of the allowed vocabulary and syntax structures and the time that will be required to train an operator, with more complicated voice protocols requiring more extensive training periods.

With regard to human eye control capabilities, our chief concerns in the design of OASIS are the precision and dynamics of voluntary eye movements and the characteristics of involuntary movements which impose noise and bias on the voluntary patterns. We are primarily concerned with conjugate translation movements of the eyes in which the two eyes move together (conjugately) to scan a flat plane (display screen) oriented perpendicular to the line-of-sight; convergent and divergent movements of the eyes will eventually be of interest when OASIS is used for interacting with three-dimensional displays (e.g., as proposed by Wixon [1983]). A useful indication of baseline noise in the eye movement system is provided by eye-tracking records of subjects who are attempting to fixate a stationary target; results from such experiments indicate that eye movement noise during steady fixation could move the eye over a range of up to one degree, so greater accuracy with an oculometer would appear to be of little value for OASIS. Other experimentally demonstrated phenomena suggest that various complex feedback situations might develop if a feedback cursor is yoked loosely to eye movements in OASIS so that cursor movement might stimulate eye movements which would produce further cursor movement and so on. Adjustments to cursor shape, color, and dynamics could probably be determined in order to avoid such problems.

There are two significant coordination issues associated with OASIS: (1) coordination between eye and voice aspects of OASIS and (2) coordination of eye and voice activities required by OASIS with other information-processing activities. Effective use of OASIS will require the operator to accurately coordinate the timing of voice commands with concurrent eye fixations or associated cursor positions. Voice time-tagging may be required to enable the operator to deliberately associate the graphic situation at a given time (i.e., eye/cursor position) with a voice command. One or more command words (e.g., "NOW", "GO", "PLEASE",...) could be used simultaneously to indicate the end of a command string, to order the execution of the command, and to indicate the time correspondence between eye and voice channels. Since the voice recognizer will take some time (say, .5 to 1.5 seconds) to recognize each time-tagged word, it is appropriate for the system to project the cursor or eye fixation position back in time a few seconds using a data buffer to compensate for the processing lag. Coordination of eye and voice actions required by OASIS with other information-processing channels and activities is a serious issue for the identification of appropriate applications for OASIS. Tasks which require extensive voice communication with other operators or extensive visual monitoring away from the display screen might generate conflicts with OASIS control actions. Manual control processes often require visual guidance and so could exhibit the same problem. Manual processes could also be very difficult to perform in conjunction with OASIS eye-voice controls because together they might overload the total attention capability of the operator. At the same time, it should be noted that because of OASIS' naturalness and efficiency, it is unlikely that another equally powerful interface mechanism could be devised which would impose a lesser demand on the operator's attention.

Some special types of fatigue may arise from extensive use of OASIS. Long-term performance of deliberate eye movement might, under some conditions, fatigue the operator's eye muscles. The requirement to keep the head within a relatively small envelope (one cubic foot with the head-tracking mirror option and one cubic inch without the mirror option) for corneal reflex point-of-regard oculometers could produce fatigue in the back and neck. Some fatigue could be reduced through good ergonomic design of the operator's seat and workstation. Attempts to achieve precise voice control in order to accommodate the limitations of voice recognition equipment could strain the vocal system.

1.2.4 Generic Characteristics of Suitable OASIS Applications

The key motivation of the OASIS control concept is the idea that direct eye-voice communication is both natural and efficient. A primary implementation of the concept is to provide an interface with a graphic display. Operator eye movements would control a screen cursor, and operator voice commands would provide ancillary discrete control. The most suitable application areas for such interfaces are determined by both operator and task characteristics.

The operator that is most benefited by OASIS is one who would otherwise be overwhelmed with manual interaction requirements. For the case of disabled individuals especially, OASIS could serve to overcome limitations in manual performance capability. This consideration also applies to individuals whose motor capabilities are severely impaired by external factors such as high G-forces or a constraining suit.

The tasks that will be most suitable for use of OASIS are those with high information-processing workload, especially in the visual and verbal modalities. The requirement to designate and manipulate spatial symbols is well-addressed by OASIS. Opportunities for visual and verbal information to be used in complementary fashion are particularly indicative of OASIS benefit. At the same time, suitable applications cannot involve high visual and verbal channel loadings beyond those deriving from use of OASIS (e.g., the operator cannot be in constant voice communication with other operators).

The system with which OASIS is used must be computer-based and should ideally contain a fair measure of intelligent capabilities. Automated voice is an excellent medium for communicating with an intelligent system (i.e., one of the artificial intelligence systems for inference, data management, etc.) because speech is the preferred medium for intelligent communication between humans. When the data being managed by the system is spatial or conducive to a spatial representation, OASIS is especially appropriate.

1.2.5 Recommended Application Areas

A broad variety of candidate OASIS applications were examined, including extravehicular activity in space, air traffic control, computer-aided design/manufacturing/engineering, computer interface for handicapped people, cartography, medical research, teleoperator control, tactical display interaction, and target acquisition. These applications are listed in Figure 1-2, along with indications of which descriptive criteria are expected to be relevant to each application. This list is by no means exhaustive, but it does represent a diverse set of problem and task areas. These areas serve as an initial point for our investigation of applications for OASIS.

1.2.6 Plan for OASIS Prototype Development

Because of the novelty of the OASIS concept, applied laboratory research must be conducted before any particular application can be definitized. Research questions which must be answered before any specific application is developed include:

- Cursor Dynamics -- How should the feedback cursor be dampened in response to the eye fixation positions received from the oculometer?

	HIGH MANUAL INTERACTION WORKLOAD	HIGH INFORMATION PROCESSING LOAD	FOCUS ON SPATIAL INFORMATION	MOTOR-IMPAIRED OPERATOR	POTENTIAL APPLICATION FOR ARTIFICIAL INTELLIGENCE
AIR TRAFFIC CONTROL	•	•	•		•
EXTRAVEHICULAR ACTIVITY IN SPACE	•		•	•	•
COMPUTER-AIDED DESIGN/ MANUFACTURING/ENGINEERING		•	•		•
COMPUTER INTERFACE FOR DISABLED PEOPLE				•	•
CARTOGRAPHY	•	•	•		•
MEDICAL ANALYSIS		•	•		•
TELEOPERATOR CONTROL	•	•	•		
TACTICAL DISPLAY INTERACTION		•	•		•
TARGET ACQUISITION	•	•	•		

Figure 1-2. OASIS Applications Classified by Benefit Criteria

- Cursor Type -- What degree of unobtrusiveness must the cursor demonstrate? That is, what is its optimum brightness, color, and shape? Should it flicker or rotate in some manner?

- Oculometer/Speech Interaction -- What is the optimum combination of eye activity and voice commands for controlling cursor movement? For example, how would user input be manipulated to position the cursor on a stationary point or on a moving target? More specifically, how can cursor movement be time-tagged to speech commands?

The following steps are necessary to build a system to answer these questions:

1. Establish a laboratory incorporating a state-of-the-art oculometer, speech recognizer, color graphics display, feedback cursor generation, and real-time data collection.

2. Define a series of generic graphics manipulation tasks, including positioning the cursor on stationary points and moving targets in both casual and time-stressed situations.

3. Define alternative methods of combining speech commands and eye fixation control to accomplish each graphics task.

4. Develop a baseline algorithm to transform eye movement data into a series of points-of-fixation, perhaps filtering out measurement noise and involuntary components of eye movements.

5. Develop a baseline algorithm for dampening the movement of the feedback cursor in response to derived eye fixation points.

6. Develop a set of subject performance measures, including completion time, accuracy, number of errors, accidental activations, and operator workload (which is loosely related to pupil diameter, an output of the oculometer).

7. Perform a series of multi-factor experiments to iteratively evaluate and refine the dampening and smoothing algorithms, the cursor characteristics, and the speech/oculometer control combination.

1.3 OVERVIEW OF PHASE II REPORT

The principal objective of this effort was to develop an eye- and voice-controlled, human-computer interface which provides a control mechanism that is natural, efficient, accurate, and low in workload. Meeting this objective involved the following four general tasks:

1. Building a laboratory facility for OASIS interface experimentation and system development.

2. Determining the characteristics of an optimal eye-voice interface.

3. Demonstrating the utility of this interface for typical applications involving tactical display interaction.

4. Preparing engineering development plans for the final OASIS design.

The remainder of this report describes our approach and activities in addressing each of the general tasks. The four tasks are considered to be project goals, and this report is a progress report on the extent of accomplishment of each goal. Overall, these goals were an ambitious undertaking, and we feel that we have met or at least made substantial headway on each of the four goals.

Section 2 presents Task 1 -- Establishing the OASIS laboratory. This goal has been accomplished exactly as planned. The OASIS laboratory facility is described in terms of both hardware and functional capability. As planned, we have successfully integrated hardware and software for the functions of eye tracking, voice recognition and interpretation, graphic display management, experimental control, and data collection/analysis. Conventional graphics devices (viz. keyboard and mouse) have been incorporated in order to provide a baseline to which OASIS performance is compared. Additionally, capabilities for efficient storage and replay of data have been implemented.

Section 3 describes Task 2 -- The OASIS experimental program. The goal of determining characteristics of an optimal OASIS interface was partially accomplished. A set of controlled experiments was performed using a single complex designation task and investigating the effects of visual feedback and filtering algorithms on OASIS performance in a factorial design. For this task, OASIS performance was compared to a conventional interface (mouse). The results showed great promise for OASIS and that more extensive research is required to fully optimize the OASIS interface.

Section 4 discusses Task 3 -- The construction of four interactive demonstration tasks which were successfully designed and implemented. These simulated tasks were developed via an innovative procedural language described in Section 2. The tasks are:

- Remote manipulation,
- Tactical targeting/firing,
- Assembly-line product inspection, and
- Multiple vehicle control.

Due to limitations of resources, controlled experiments comparing OASIS versus conventional interface devices were not performed on these demonstration tasks.

Task 4 -- Preparation of engineering development plans -- was not reached for two important reasons. First, the research described in Section 4 showed that more generic research is required before an optimal interface can be specified in detail. Second, a strong result of those experiments is that optimal features are very application-dependent, and we have not been committed to a specific application.

Section 5 presents our different, ongoing approaches to "next steps" for OASIS development. These approaches include:

- Updated listing of potential application areas for both the government and private sector,
- Funded OASIS-related research at Analytics,
- Submitted concept papers which describe example OASIS interface designs,
- Listing of OASIS presentations made, and
- Listing of OASIS magazine/newspaper articles.

2. TASK 1 -- ESTABLISH OASIS LABORATORY

2.1 OASIS LABORATORY CONFIGURATION

The OASIS Testbed Laboratory was developed to provide a realistic task environment for the exploration and development of eye-voice interface technology. Figure 2-1 shows the layout of the principal functional components of the OASIS testbed (development tools such as printer, tape recorder, and video camera are not shown). The test subject/user wears a headset microphone and is seated in an adjustable chair with headrest, facing a keyboard, mouse, and two color terminals. The larger of the terminals is located directly in front of the subject and is used to present the runtime task display. The smaller is placed slightly to the side of the larger and is used to present text material when the user is enrolling the recognition vocabulary. A small mirror is located just in front of the larger display terminal to deflect the IR beam from the optical head into the user's right eye. The experimenter controls experiment runs from his own control station (display and keyboard) located to the subject's left. Oculometer calibration requires that the experimenter also have access to the ocular subsystem located to the subject's right.

The OASIS testbed hardware configuration is presented in Figure 2-2. The same testbed components are shown here as in the preceding figure, and in the same arrangement. OASIS testbed software resides on the Masscomp MCS-531, a 68000-based Unix machine. Testbed graphics are presented on a 19-inch color raster display operating at 60 Hz, non-interlaced. The Applied Science Laboratories Model 1996 Eye View Monitor System (oculometer without the head-tracking option) transmits x and y eye position data and pupil diameter to the Masscomp computer 60 times a second. Data from the optical mouse is transmitted as required by movement of the mouse. Testbed communications software transforms mouse input into a 60 frames-per-second signal for use by the rest of the

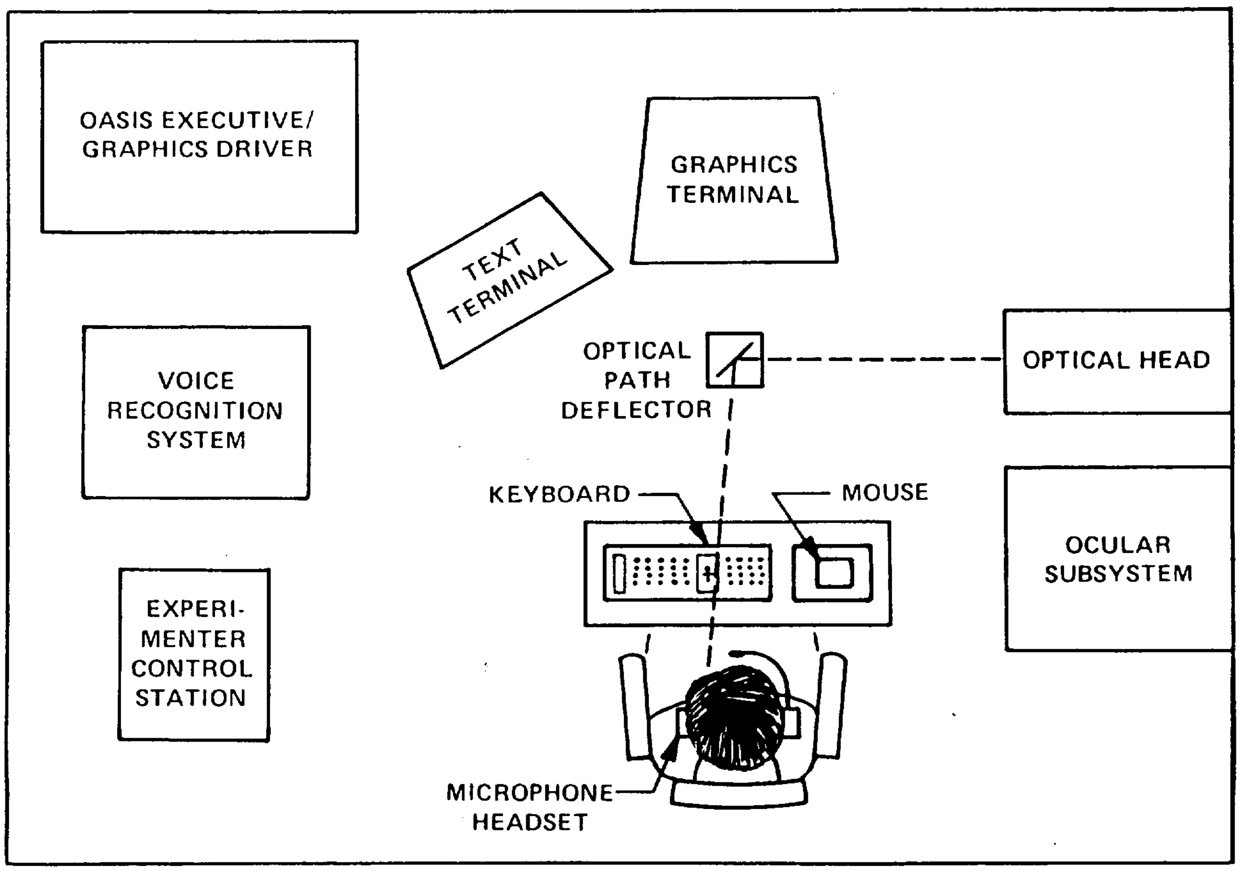


Figure 2-1. Testbed Lab Layout

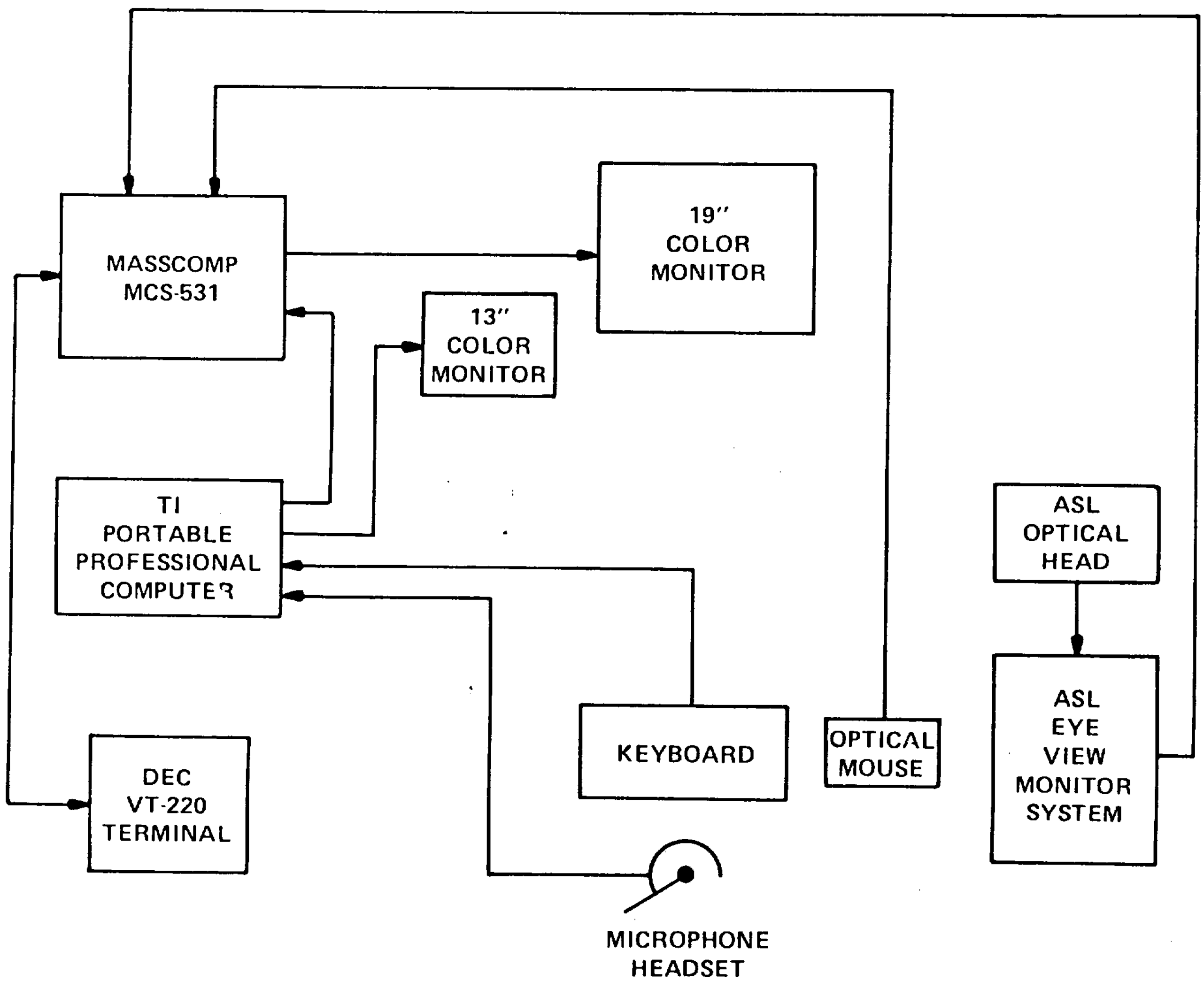


Figure 2-2. Testbed Hardware Configuration

86-21-07-FR

system. Voice recognition is accomplished by means of Texas Instruments (TI) voice hardware. Voice and keyboard inputs are passed to the Masscomp through the TI Portable Professional Computer and cannot be directly distinguished by software running on the Masscomp. Runtime control is exercised through a DEC VT-220 terminal which functions as the Masscomp's system console.

2.2 MAJOR SOFTWARE MODULES

The OASIS testbed includes four major software modules, each of which is executed as an independent program:

- icon editor
- overlay editor
- procedure compiler
- experiment runtime

Only the experiment runtime module is actually executed during an experiment or demonstration run. The other three modules are used by the experimenter to construct a task environment for use in such a run.

All OASIS testbed experimental tasks consist of exchanges between a human user and a computer-generated graphics image. As this image must be specified by the experimenter, interactive graphics tools are required. The icon editor allows the experimenter to design 16 x 16 bit icon maps -- pixel-level descriptions of task symbology. Once constructed, the icons can be embedded in static background views or manipulated as dynamic symbols at runtime. Background views are themselves constructed using the overlay editor. This program allows the experimenter to design full-screen static images in one or more of the available graphics memory planes. If two or more planes are used, multi-color overlays can be produced. At runtime, plane assignments can be made such that symbols will appear to move in front of or behind a particular overlay image.

In order to provide a flexible and realistic experimental task environment, the dynamics of the graphics image must support three separate functions:

- task simulation
- interaction with subject/user
- control by experimenter

Image dynamics during experimental runs are managed by means of programs written in a procedural language called PLEX, developed explicitly for this purpose. The program is written by the experimenter as a series of action calls grouped into labeled procedures. This program is converted by the procedure compiler into a form usable by the experiment runtime module. (See Section 4 for additional details regarding PLEX.) During runtime, the program provides the control necessary to advance the simulation and to allow for interaction between user and simulation as well as for control by the experimenter.

The experiment runtime module reads in a PLEX program specified by the experimenter, loads the icons and overlays referenced by that program, and executes the program's initial procedure. From this point on, the runtime module loops through the following steps 60 times a second:

1. Retrieve continuous input (that is, eye or mouse data).
2. Retrieve discrete input (that is, voice or keyboard data).
3. Execute procedures driven by discrete input.
4. Execute clock-driven procedures.
5. Update the position of graphics objects based on continuous input and internal models and filters.
6. Build and display the new graphics image.

At the end of an experimental run, the raw data retrieved from both continuous and discrete input devices can be dumped to a disk file. If the original PLEX program is also saved on disk, this data can be read in by the experiment module and reprocessed. With the exception of the data retrieval functions, the difference between live and recorded data is transparent to the experiment run-time module.

When an experimental run is replayed, several options are available to the experimenter. First, the data can be replayed in demonstration mode. In this mode, the replay takes place in real time and the dynamics of the screen image are identical with the original run. Second, the experimenter can replay the data in stop-action mode. In this mode, the replay can be stepped forward one frame at a time or arbitrarily advanced to any position in the recording. While in stop-action mode, the experimenter control station continuously displays the current frame number and identifies each PLEX procedure as it is executed. Third, the data can be replayed in data-extraction mode to build data files for post-analysis. In this mode, the experimenter may select the entire file or some frame window within the file for replay, as well as those data items which are to be extracted. As the replay takes place, raw data, the position of graphics objects, or the distances between them are dumped to files for subsequent analysis. No graphics image is generated for viewing while the data-extraction replay is taking place.

3. TASK 2 -- CONDUCT GENERIC TASK EXPERIMENTS

The second goal of the Phase II research was to develop generic experimental tasks using the software testbed described in Section 2. Initial experiments would assess the feasibility of the overall OASIS concept and determine the relative roles of some of the key components of the system. Future experiments would be used to iteratively evaluate alternative OASIS processing algorithms and visual feedback configurations in order to optimize the interface for specific applications and tasks. The results of initial experimentation are reported in this section.

3.1 SELECTION OF THE INDEPENDENT EXPERIMENTAL FACTORS

The basic system components of OASIS are an eye movement measurement device (oculometer), a voice recognition/interpretation mechanism, and a central processor to combine the eye and voice inputs into system commands. However, OASIS is not a simple linking of these components. There are a number of substantive and complex human factors issues that must be addressed before a new system such as OASIS can be compared with other forms of human-computer interface (e.g., keyboard, mouse). Some of the most important of these issues are:

- Determining visual attention from eye position data,
- Providing the operator with feedback which is both informative and not distracting,
- Recognizing and interpreting voice inputs, and
- Coordinating the voice and ocular inputs.

The initial OASIS experiments addressed the first two of these issues -- deriving visual attention and presenting visual feedback.

3.1.1 Determining Visual Attention

The eye, when looking at something, is in constant motion. Intentional movements are saccadic -- rapid conjugate movements by which we change fixation from one point to another, for example, in scanning a visual scene or reading. There are various types of partially voluntary or involuntary eye movements, including slow and smooth pursuit-tracking movements, smooth compensatory movements, vergence movements, blinks, and a variety of low-amplitude movements observed during attempted steady fixation of a point (Young and Sheena, 1975; Eizenman et al., 1984). Because of these various and everpresent motions, the momentary point-of-gaze, as computed by an oculometer, is only an approximation of the visual attention point.

Ideally, visual attention would simply be the result of filtering out all involuntary movements or noise. In practice, it is not usually obvious whether a change in eye position from one frame to the next (our system operates at 60 Hz) reflects attention or noise. Also, there are a number of methods (algorithms) for filtering noisy data. For this preliminary study, we chose three different levels of eye movement processing or filtering, each preceded by a correction which eliminates blinks:

1. Raw (R) condition: The momentary point-of-gaze output from the oculometer drives the system. This is the control condition.
2. Smooth (S) condition: Visual attention is the running average of the last 15 frames of data (at 60 Hz, this is 250 msecs). This very simple algorithm just smooths, but does not distinguish between voluntary and involuntary movements.
3. Eye Point Tracker (EPT) condition: A second-order Kalman filter -- which is a weighted sum of the current and past history of both eye position and velocity -- is applied to the raw data in addition to a correction for voluntary movements. This condition represents the most complex level of processing.

We use the term "computed attention" to refer to where the system 'thinks' the operator's visual attention is at any given point in time. In the R condition, computed attention is just the oculometer output with blinks removed. In the two filtering conditions, computed attention is the processed oculometer output.

3.1.2 Visual Feedback

Visual feedback is another important and complex issue. It is intuitively compelling that, in a largely visual task, performance could be enhanced with information as to where the system thought the user was looking. On the other hand, if such information were very salient, the operator might pay more attention to it than to what he was tasked to look at, thus degrading performance.

For this study, we have selected three disparate visual feedback conditions:

1. No feedback (NF) condition: The system provides no visual feedback of computed attention. This is the control condition.
2. Discrete (D) condition: The system provides a binary type of feedback. No feedback is provided to the subject until computed attention first falls within a criterial distance from the target; at this point, the target changes color.
3. Continuous (C) condition: The system provides continuous and complete visual feedback in the form of a feedback cursor which represents computed attention at each moment of time.

3.2 SELECTION OF THE EXPERIMENTAL TASK AND DEPENDENT VARIABLES

There were a number of types of experimental tasks considered, including:

- Simple designation (subject moves his gaze to some point of fixation specified by the experimenter),
- Manipulation (subject designates a point, then performs some control action at that point),
- Simple tracking (subject follows a single moving object), and
- Complex tracking (subject tracks several objects in parallel).

In this preliminary study, no more than one type could be selected due to resource limitations. The task had to be generic, challenging, clearly measurable, and amenable to the comparison between OASIS and a conventional, manual man-machine interface. These constraints led to the choice of a designation task in which fixation (target) points appear at random locations on a display, and the subject must move his/her gaze to that point as quickly as possible and keep his/her gaze on the target point for some (brief) specified period of time.

It is desirable to look at measures of both speed and accuracy. Though a number of performance measures have been examined during the course of OASIS research, we have selected a single speed measure and a single accuracy measure to report on:

1. Time to Acquire Target (TTAT) is the time elapsed from the time a target appears to the time when computed attention first falls within a criterial distance from the target point. TTAT is a speed measure.

2. Time to Stabilize on the Target (TTST) is the time elapsed from the time of acquisition of the target (as defined in 1.) to the time when computed attention falls within the criterial distance from the target for 30 consecutive frames (.5 sec). TTST is an accuracy/stability measure.

3.3 EXPERIMENTAL METHOD

Six subjects, three males and three females, were calibrated on an ASL 9600 infrared tracking oculometer in a dimly lit room with incandescent illumination. Subjects were familiarized with the system by practicing a simple graphics designation task, alternatively using the mouse and OASIS as the pointing device. The graphics tasks were presented on a high-resolution display (800 x 600 pixels) which was positioned at a 42-inch viewing distance and occupied a field-of-view of 17.3 degrees horizontally and 12.5 degrees vertically. Subjects practiced the training task until they met a time criterion. Subjects then began experimental trials. A list of the conditions, along with a condition code, is presented in Table 3-1.

Table 3-1. Experimental and Control Conditions

Experimental Conditions	Code
Continuous Feedback/Raw Eye Filter	CR
Continuous Feedback/EPT Eye Filter	CE
Continuous Feedback/Smooth Eye Filter	CS
Discrete Feedback/Raw Eye Filter	DR
Discrete Feedback/EPT Eye Filter	DE
Discrete Feedback/Smooth Eye Filter	DS
Control Conditions	
No Feedback/Raw Eye Filter	NFR
No Feedback/EPT Eye Filter	NFE
No Feedback/Smooth Eye Filter	NFS
Mouse	M

The mouse control condition was presented first followed by the no feedback (NF) condition. The subjects were only presented with one NF condition even though three NF conditions are listed in Table 3-1. The three NF control conditions were subsequently created from the single condition by replaying the eye movement history through three eye-filtering algorithms -- raw (NFR), EPT (NFE), and smooth (NFS). The six experimental conditions were then presented to the subjects, counterbalancing the ordering of the conditions for each subject. Oculometer calibration accuracy was checked at pre- and post-experimental periods to ensure that calibration quality was maintained throughout the experiment.

For each condition, the subject performed a simple search task. The display presented a field of 27 randomly located blue dots, each subtending a half-degree visual angle. The dots were overlayed on a light blue background. Every two seconds, a dot was targeted with a color change from blue to black. The subject's task was to point to the black target with his/her eye during oculometer trials and to point to the target with the mouse during mouse trials. Eighteen targets were presented for each condition and trial. Each trial was 36 seconds.

For the continuous feedback trials (CR, CE, and CS), a cursor continuously displayed the OASIS computed attention point. For the discrete feedback trials (DR, DE, and DS), the subject was only provided with a binary indicator of a successful target acquisition and stabilization. If the currently targeted dot turned green, then the OASIS computed attention was within a one-degree visual envelope of the true target position.

3.4 RESULTS

Data analyses included raw data summaries, an analysis of variance (ANOVA) of the experimental factors, and an error analysis. Section 3.4.1 presents detailed raw data summaries for each OASIS and mouse condition in terms of Time to Acquire Target (TTAT) and Time to Stabilize on the Target (TTST).

ANOVAs for TTAT and TTST are presented in Sections 3.4.2 and 3.4.3. Nine conditions were input into the ANOVA, representing three feedback levels (discrete, continuous, and no feedback) and three filter levels (raw, smooth, and EPT). The ANOVAs were performed with the following modifications to the raw data set:

- All TTAT and TTST time data was converted to speed measures (that is, by taking the inverse of the raw time value).
- Trials with missing data (that is, trials where the subject had failed to acquire or failed to stabilize the target) were assigned raw time values of 10 seconds.

Error analyses for failures to acquire and stabilize targets are presented in Sections 3.4.4 and 3.4.5.

3.4.1 OASIS Versus Mouse -- Raw Data

Table 3-2 presents TTAT means and standard deviations for each OASIS and mouse condition. The raw TTAT scores were averaged across all subjects for cases where target acquisition was successful (1063 of 1080 cases). Recall that successful acquisition is the ability of the OASIS computed attention cursor or the mouse cursor to fall within a one-degree accuracy envelope of the target before the target expires.

Table 3-2. TTAT Means and Standard Deviations

	Mean	Standard Deviation
No Feedback/Raw Eye Filter	0.49	0.091
No Feedback/EPT Eye Filter	0.57	0.194
No Feedback/Smooth Eye Filter	0.68	0.092
Continuous Feedback/Raw Eye Filter	0.56	0.189
Continuous Feedback/EPT Eye Filter	0.57	0.179
Continuous Feedback/Smooth Eye Filter	0.72	0.109
Discrete Feedback/Raw Eye Filter	0.54	0.134
Discrete Feedback/EPT Eye Filter	0.60	0.179
Discrete Feedback/Smooth Eye Filter	0.74	0.189
Mouse	0.86	0.195

Every oculometer condition outperformed the mouse in time to acquire targets. The mouse lagged the slowest OASIS condition by at least a tenth of a second and lagged the fastest OASIS condition by almost four tenths of a second. Of the OASIS conditions, the raw eye filter showed the best performance, closely followed by EPT, with the worst performance by the smooth filter. This result is expected as the smooth filter exhibits the greatest cursor dampening effect; the smooth filter is a running average of the last 15 frames. On the other hand, the EPT filter eliminates old frame history whenever a major saccade is detected.

Table 3-3 presents TTST means and standard deviations for each OASIS and mouse condition. The raw TTST scores were averaged across all subjects for cases where target stabilization was successful (983 of 1080 cases). Recall that a perfect TTST score is the time elapsed from the time of acquisition of the target to the time when computed attention falls within a one-degree visual angle accuracy envelope for 30 consecutive frames or 0.50 seconds.

Table 3-3. TTST Means and Standard Deviations

	Mean	Standard Deviation
No Feedback/Raw Eye Filter	0.63	0.238
No Feedback/EPT Eye Filter	0.54	0.133
No Feedback/Smooth Eye Filter	0.53	0.100
Continuous Feedback/Raw Eye Filter	0.66	0.285
Continuous Feedback/EPT Eye Filter	0.54	0.173
Continuous Feedback/Smooth Eye Filter	0.54	0.128
Discrete Feedback/Raw Eye Filter	0.63	0.238
Discrete Feedback/EPT Eye Filter	0.54	0.177
Discrete Feedback/Smooth Eye Filter	0.55	0.152
Mouse	0.52	0.088

As expected, the mouse exhibited an almost perfect stabilization score (0.52). However, the OASIS EPT and smooth filtering conditions also exhibited an almost perfect stabilization score (0.54 and 0.55).

3.4.2 Processed TTAT -- Time to Acquire Target

A repeated measures ANOVA was performed with missing TTAT data set to 10 seconds before conversion of all TTAT scores to inverse time or speed. The raw value of 10 was selected so that after conversion the speed value of the failed acquisitions approached zero. The missing value of 10 was selected as the raw outlier score since the maximum value for successful acquisitions was two seconds or the target duration time.

An ANOVA was performed with two trial and two grouping factors. The trial factors were feedback mode and eye filter mode, with feedback at three levels (continuous, discrete, no feedback) and eye filter also at three levels (raw, smooth, EPT). The grouping factors were target distance and subjects.

For the feedback and eye filter trial factors, Huynh-Feldt probabilities showed significant main effects ($F(2,170) = 11.33$, H-F $p < 0.00005$ and $F(2,170) = 295.05$, H-F $p < 0.00005$, respectively). The two-way interaction of feedback and filter was also significant ($F(4,340) = 2.67$, H-F $p = 0.0320$). The grouping factors, subject and target distance, also showed significant main effects ($F(5,85) = 7.14$, H-F $p < 0.00005$ and $F(17,85) = 2.32$, H-F $p = 0.0059$, respectively). However, when looking at two-way interactions between trial and grouping factors, only subject grouping was significant for both the feedback and filter factors ($F(10,170) = 5.73$, H-F $p < 0.00005$ and $F(10,170) = 5.01$, H-F $p < 0.00005$, respectively). Target distance did interact with feedback mode ($F(34,170) = 1.83$, H-F $p = 0.0068$).

Figure 3-1 presents average TTAT performance scores as inverse time (or speed measures) for each feedback and eye algorithm combination, including the speed values assigned to unsuccessful acquisitions. The results indicate that the raw eye filter is fastest and that smooth is slowest. This is not surprising as dampening slows movement. However, EPT performance was almost as good as the raw eye filter. Even though the EPT filter uses dampening to eliminate the noise of microsaccades, the EPT filter ignores all eye track history following a major saccade. Concerning feedback modes, no feedback was the fastest; continuous feedback was the slowest. This confirms the expectation that the cursor might disrupt performance by leading eye fixation. However, this effect is small.

Figure 3-1 also plots the average mouse speed for comparison to the OASIS conditions. The result indicates that the OASIS interface has an advantage over conventional graphic controllers in terms of target acquisition speed. As reported in Section 3.4.1, considering the average raw time to acquire targets, the mouse was a tenth slower than the slowest eye filter (smooth) and almost four tenths slower than the faster eye filters (EPT and raw).

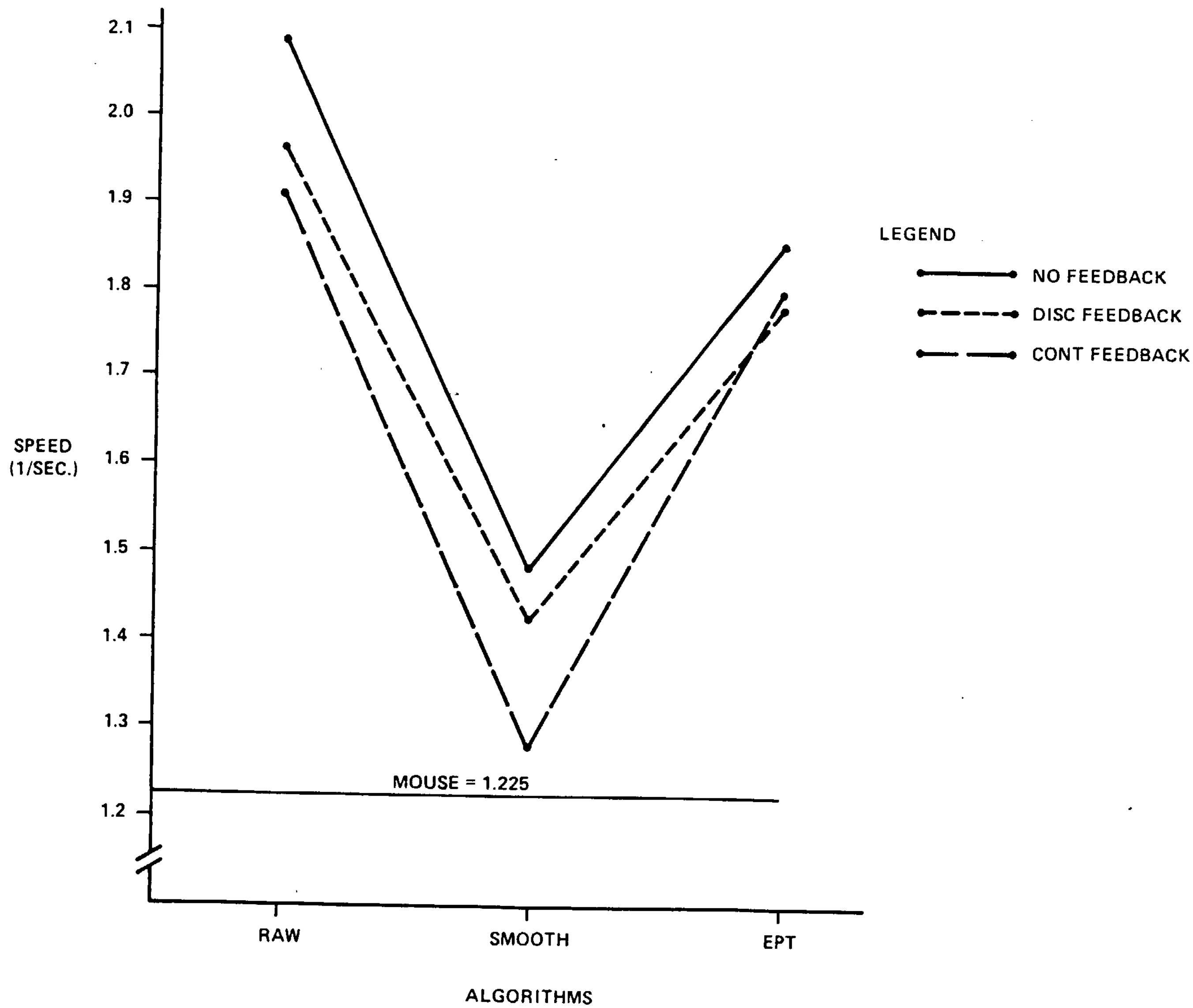


Figure 3-1. Average TTAT Scores by Feedback and Filter

3.4.3 Processed TTST -- Time to Stabilize on the Target

A repeated measures ANOVA was performed for TTST measures similar to that described for TTAT in Section 3.4.2. Again, for cases where stabilization was unsuccessful, a value of 10 seconds was assigned to the cell before conversion of all data to speed scores. The same grouping and trial factors were also repeated.

For the feedback and eye filter trial factors, Huynh-Feldt probability tests found significant main effects ($F(2,170) = 8.39$, H-F $p = 0.0003$ and $F(2,170) = 37.44$, H-F $p < 0.00005$, respectively). The two-way interaction of feedback and filter was not significant.

Only the grouping factor of subject showed a significant main effect ($F(5,85) = 7.48$, H-F $p < 0.00005$). Target distance was not significant for TTST. When looking at two-way interactions between trial and grouping factors, subject grouping was again significant for both the feedback and filter factors ($F(10,170) = 8.14$, H-F $p < 0.00005$ and $F(10,170) = 1.86$, H-F $p = 0.0542$, respectively). Target distance did not interact with the filter factor, but was almost significant with the feedback factor ($F(34,170) = 0.65661$, H-F $p = 0.0579$).

Figure 3-2 presents average TTST performance scores as inverse time (or speed measures) for each OASIS feedback and eye algorithm combination and for the mouse. The results for TTST are very similar to those of TTAT, considering OASIS feedback. That is, no feedback is best; continuous feedback is worst. This again confirms the suspicion that continuous feedback disrupts the task by leading the eye fixation. This decrement in performance is even more severe for TTST than TTAT.

Figure 3-2 shows that the mouse TTST is slightly better than any OASIS condition. Algorithm optimization is needed to improve stabilization times to the levels attained by a mouse. As presented in Section 3.4.1, the mouse raw

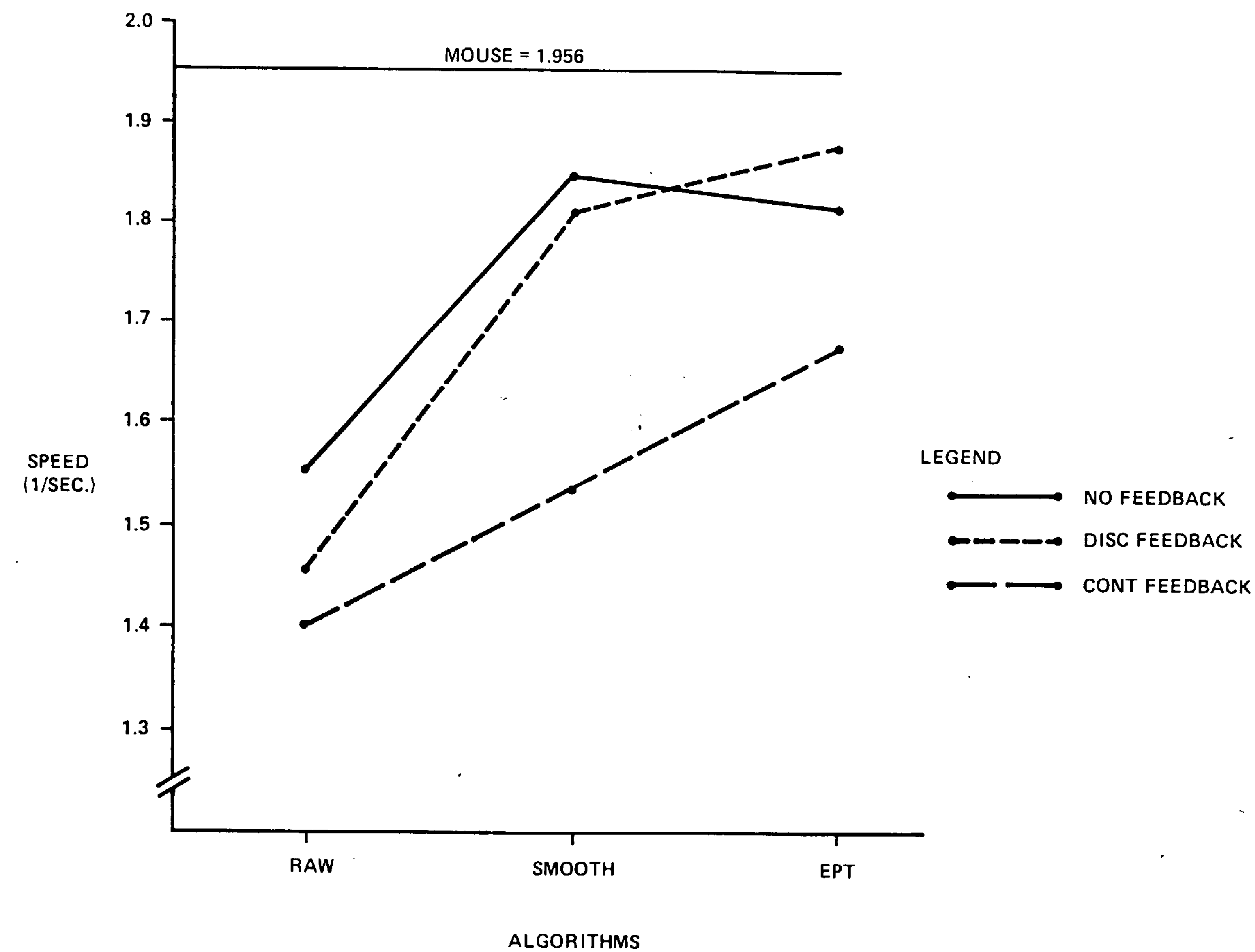


Figure 3-2. Average TTST Scores by Feedback and Filter

data was virtually perfect for stability. However, the smooth and EPT filters were nearly as good (within a tenth of a second).

The results also show that the eye filters differentially impact the TTAT and TTST performance measures. As stated in Section 3.4.1, raw eye was superior to any other filter level for acquisition time. However, when stability is the primary concern, the raw eye filter is the worst performer. EPT and smooth are considerably better. The combined results for TTAT and TTST suggest that, for the experimental task, rapid acquisition (with minimal eye filtering) must be traded off with enhanced stability (with high levels of eye filtering). Considering TTAT and TTST performance measures, the EPT is the filter of choice as targets can be rapidly acquired (almost as fast as the raw eye filter) and quickly stabilized (almost as fast as the mouse). Furthermore, it appears that continuous feedback should be avoided in favor of none or, when necessary, discrete, intermittent, or on-demand feedback.

3.4.4 Failure to Acquire Targets

Figure 3-3 presents the frequency of failures to acquire targets for each OASIS condition. For the OASIS conditions, a large number of errors occurred during continuous feedback. Continuous feedback using the smooth filter appears to be particularly distracting to the user as it more severely lags time eye position by its nature when major saccades occur.

3.4.5 Failure to Stabilize Targets

Figure 3-4 presents the frequency of failures to stabilize targets for each OASIS condition. The OASIS errors show a similar pattern to the TTST results, with the most errors occurring during conditions with the raw eye filter and/or continuous feedback.

The absolute numbers of OASIS stabilization errors shown in Figure 3-4 represent a real problem since there are only 108 observations per condition. Up to 20 percent error rates were found with the worst performance with the

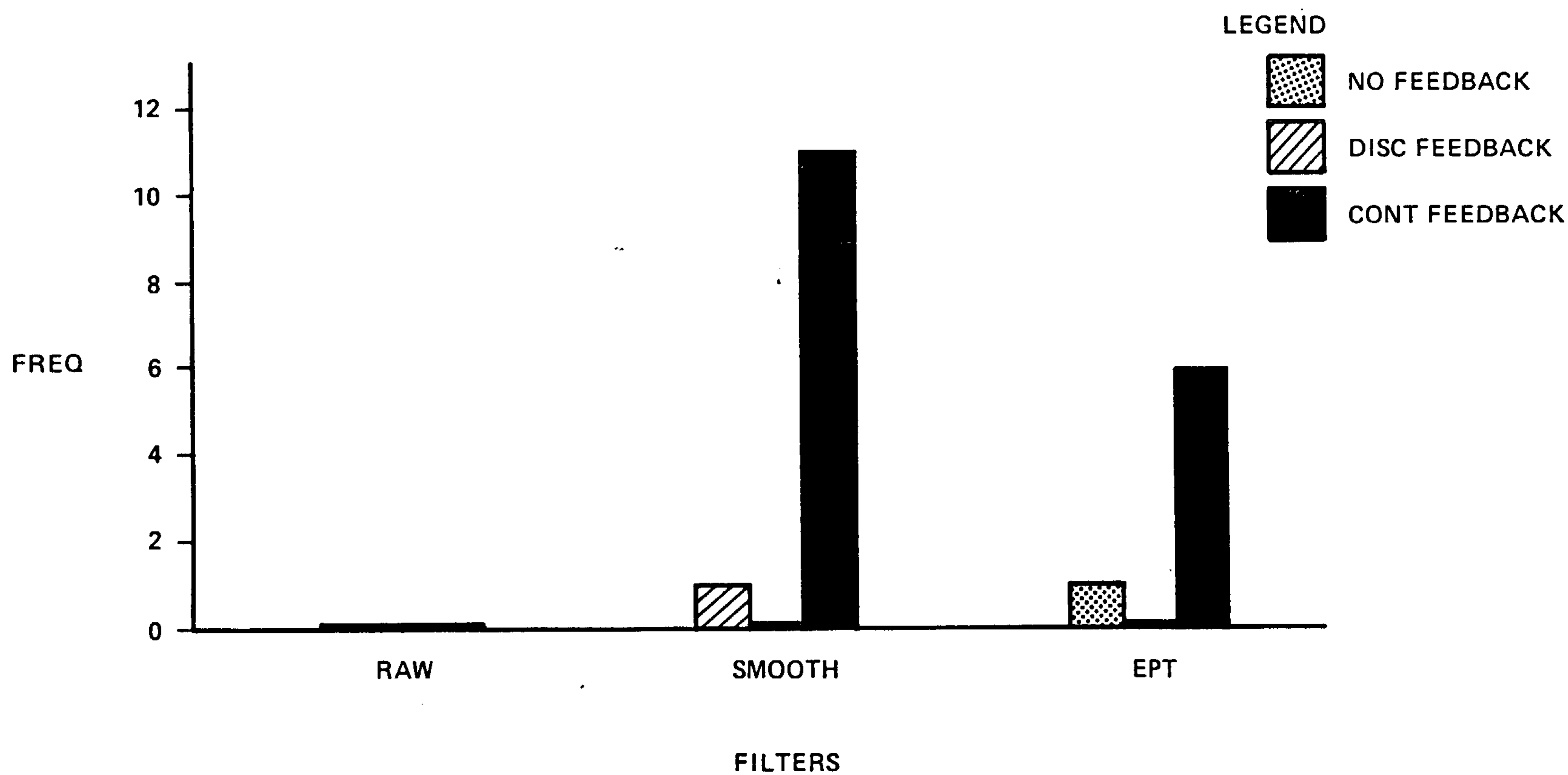


Figure 3-3. Frequency of Failure to Acquire by Feedback and Filter

86-21-07-FR

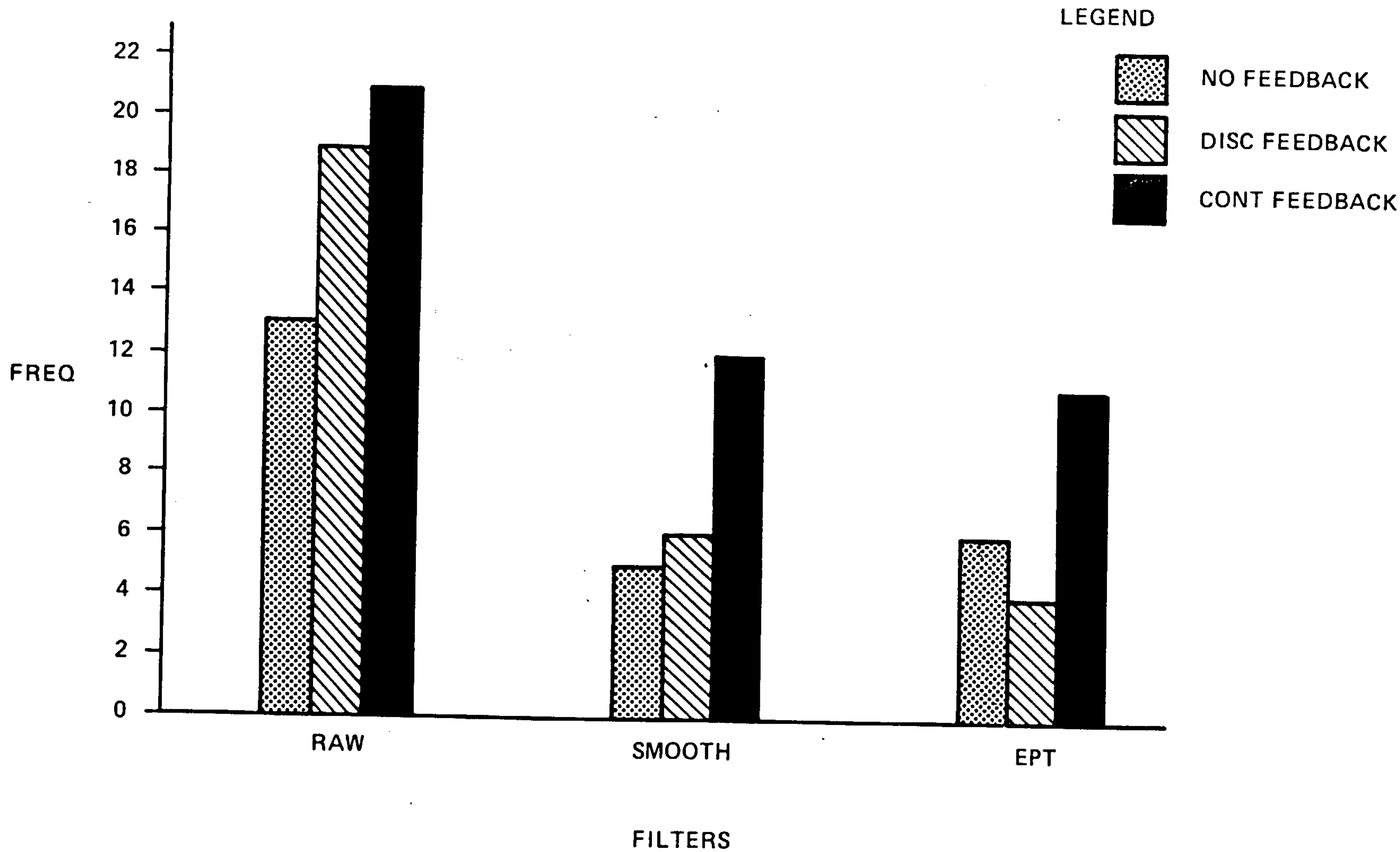


Figure 3-4. Frequency of Failure to Stabilize by Feedback and Filter

86-21-07-FR

raw-eye, continuous-feedback condition. However, the stability criterion is severe. No excursions out of a one-degree visual envelope were allowed for a full half-second, and target trials were only two seconds in length. More tolerable error rates (approximately 5 percent) were found for the smooth and EPT filters with discrete or no feedback. The overall stabilization error was 10 percent for all OASIS conditions. The need for further optimization is evident.

3.5 CONCLUSIONS

The initial OASIS prototype is a system integrating off-the-shelf components and utilizing an unoptimized set of algorithm parameters. Nevertheless, OASIS compared favorably with the conventional rapid-pointing interface. Every OASIS condition, no matter what the feedback or filtering level, outperformed the mouse in terms of target acquisition time. Of course, stabilizing on a stationary target after acquisition was a trivial task for the mouse. OASIS stabilization times were close to the mouse (on the average, within a tenth of a second) for successful stabilizations.

Further research is needed to determine the factors which contributed to the OASIS stabilization error rates. These errors may have been caused by temporary losses in oculometer calibration, eye nystagmus, or feedback/filter interactions. The results indicate that individual differences are also a contributing factor. Indeed, 37 percent of all stabilization errors were associated with a single subject. Also, subject training is likely to be a contributing factor; the most experienced subject had the smallest error rate (3 percent). Furthermore, subjects varied on their preference for specific OASIS feedback levels and exhibited their best performance under varying OASIS feedback/filtering conditions. Experiments must be performed with a much larger subject sample and over a longer time period to identify individual difference effects and to establish repeatability of results.

The estimated cost to optimize OASIS algorithms and parameters is also dependent on specific application demands. To date, algorithm development has proceeded assuming that the controlling system has no knowledge of target locations other than what is being sensed by the human operator's eyes. Depending on the application, this may or may not be a realistic assumption. If data is available on target position from alternative sensors or automatic target trackers, the OASIS feedback cursor could lock onto or capture the target closest to the operator's visual attention. The OASIS cursor would then behave in a manner very much like a mouse (that is, following a rapid excursion, the cursor would quickly stabilize on a target).

In the process of developing the laboratory facility and the sample demonstrations, it became apparent that the optimum set of eye-processing parameters for a simple designation task (such as the teleoperation task) differed from that needed for a tracking task (helicopter fire control task). Therefore, the results of the initial OASIS experiment can only be generalized to designating stationary objects. Other experiments must be designed for alternate purposes. Ideally, further experimentation will proceed after the selection of a specific application so that experimental factors can be tailored to the demands of the application.

4. TASK 3 -- DEVELOP DEMONSTRATION APPLICATION

4.1 PLEX PROGRAMS

All experiments and demonstration applications which run on the OASIS testbed are based on programs written in PLEX. In turn, the execution of PLEX programs is based on capabilities built into the experiment runtime module. In all, over 30 different functions can be coded into PLEX procedures for runtime execution. These may be summarized under the following categories:

- set parameter(s) of a dynamic graphics object (e.g., icon, visibility, color);
- connect object to motion model or continuous data channel (oculometer or mouse);
- test object-to-object distance (pick function);
- execute or conditionally execute a procedure;
- assign procedure execution to discrete data entry;
- assign procedure execution to a clock time.

When combined into procedures, these primitive functions can provide considerable power and flexibility in simulating typical computer displays and human-computer interactions. The following pseudo-code example presents the general technique of procedural construction used in PLEX programs:

```
Procedure #1:

     set object A to display of cross-hair icon
     set object A to be visible
     set object A to be red
     connect object A to a fixed position at
          screen center coordinates
     assign Procedure #2 to execute after 120 frames
     assign Procedure #3 to execute upon input of X
```

Procedure #2:

```
set object A to display of filled-circle icon
set object A to be blue
connect object A to a linear motion model
    specified to move from left to right
    at 2 pixels per frame
```

Procedure #3:

```
terminate run
```

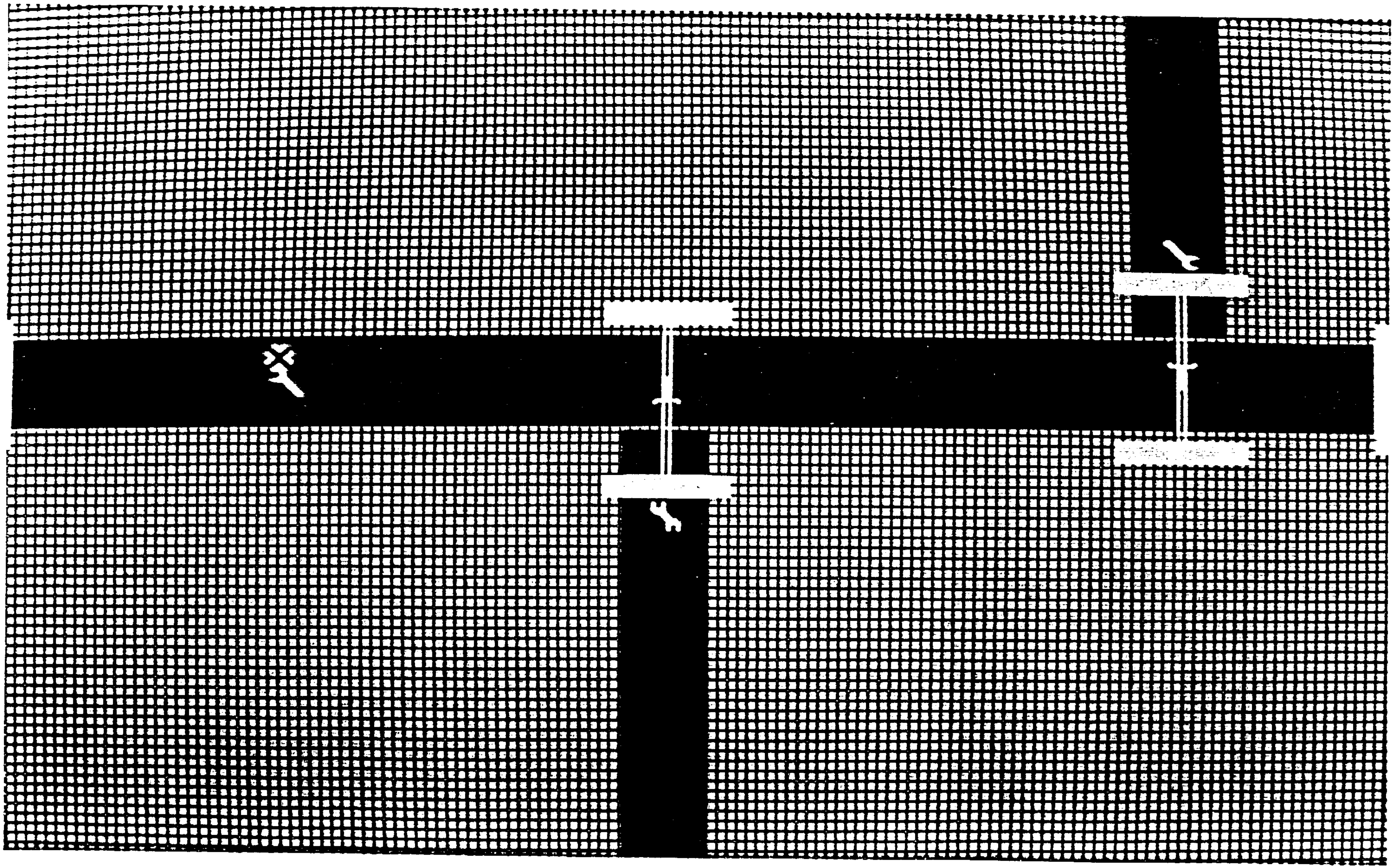
At runtime, the above procedures would result in the following scenario. A stationary red cross-hair symbol would appear at the center of the display. After two seconds, the red cross hair would be replaced (at the same location) by a blue dot which would immediately begin to move across the screen. If the subject/user never entered the X from keyboard or voice, the dot would disappear off the edge of the screen and the display would remain blank for the duration of the run. If the subject/user entered the X command, the experimental run would terminate immediately.

4.2 OASIS DEMONSTRATION APPLICATIONS

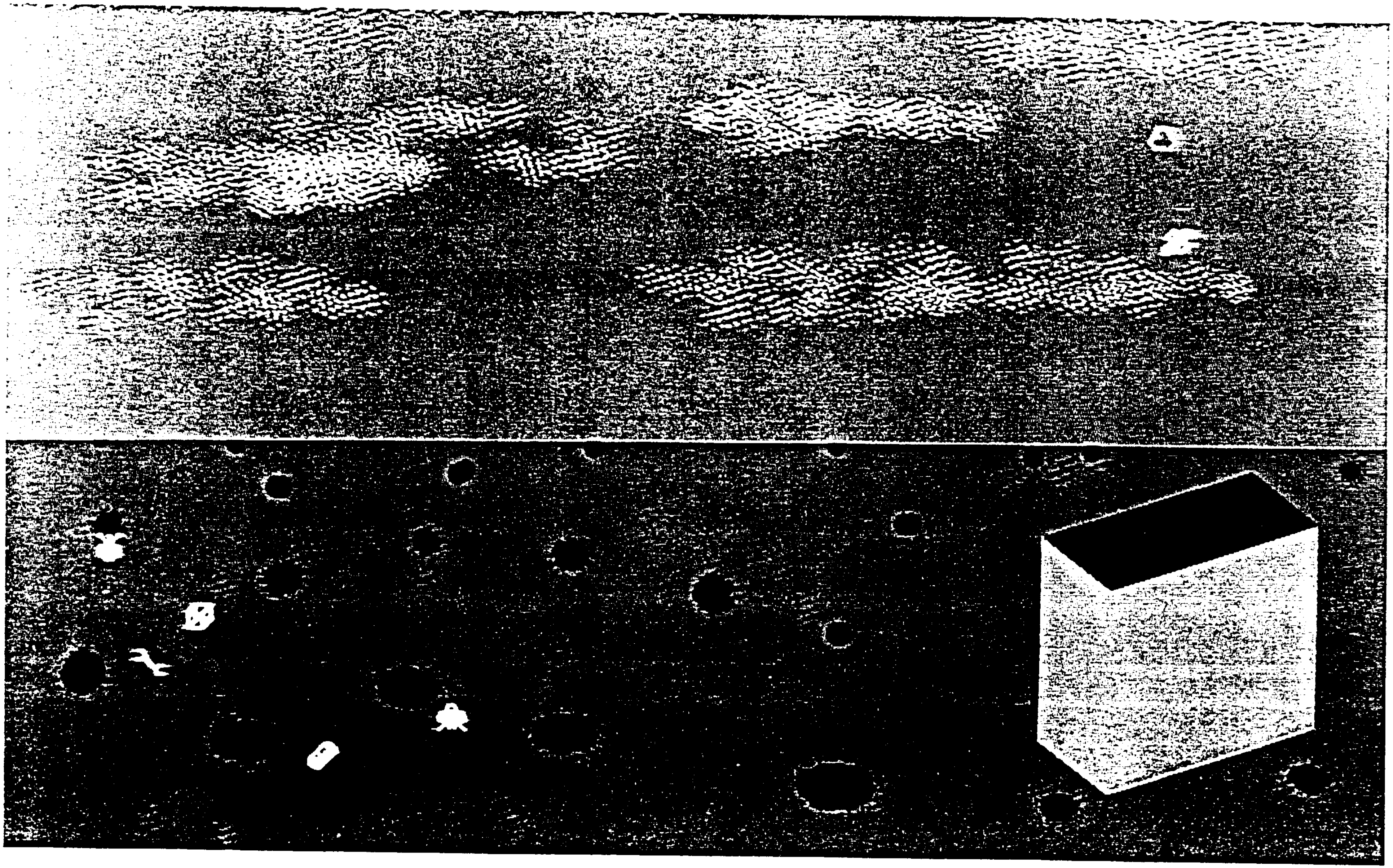
Four OASIS demonstration applications have been developed under PLEX:

- quality-control inspection
  (identify and designate moving objects)
- teleoperation of robots
  (designate and reposition stationary objects)
- air-air or ground-air fire control
  (track and designate moving objects)
- simultaneous control of six vehicles
  (monitor and control six moving objects)

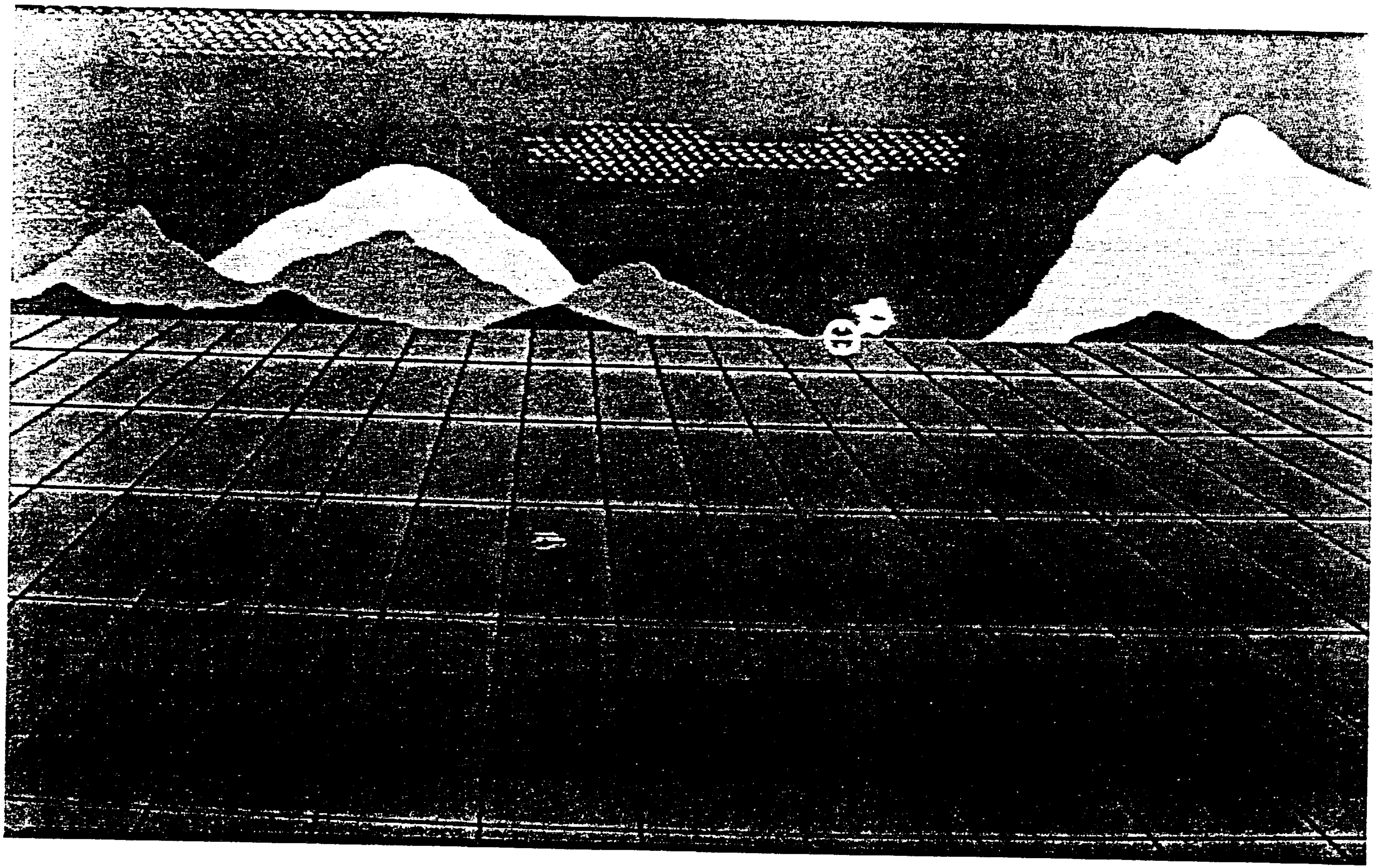
These represent a broad range of systems where eye-voice interfacing could enhance overall system performance. Figures 4-1 through 4-4 are photographs of the graphics screen images produced during actual runs of the demonstration applications. The views shown were not developed to stand as realistic scenes of an external world, but rather as typically schematic computer displays incorporating even more abstract interface symbology.



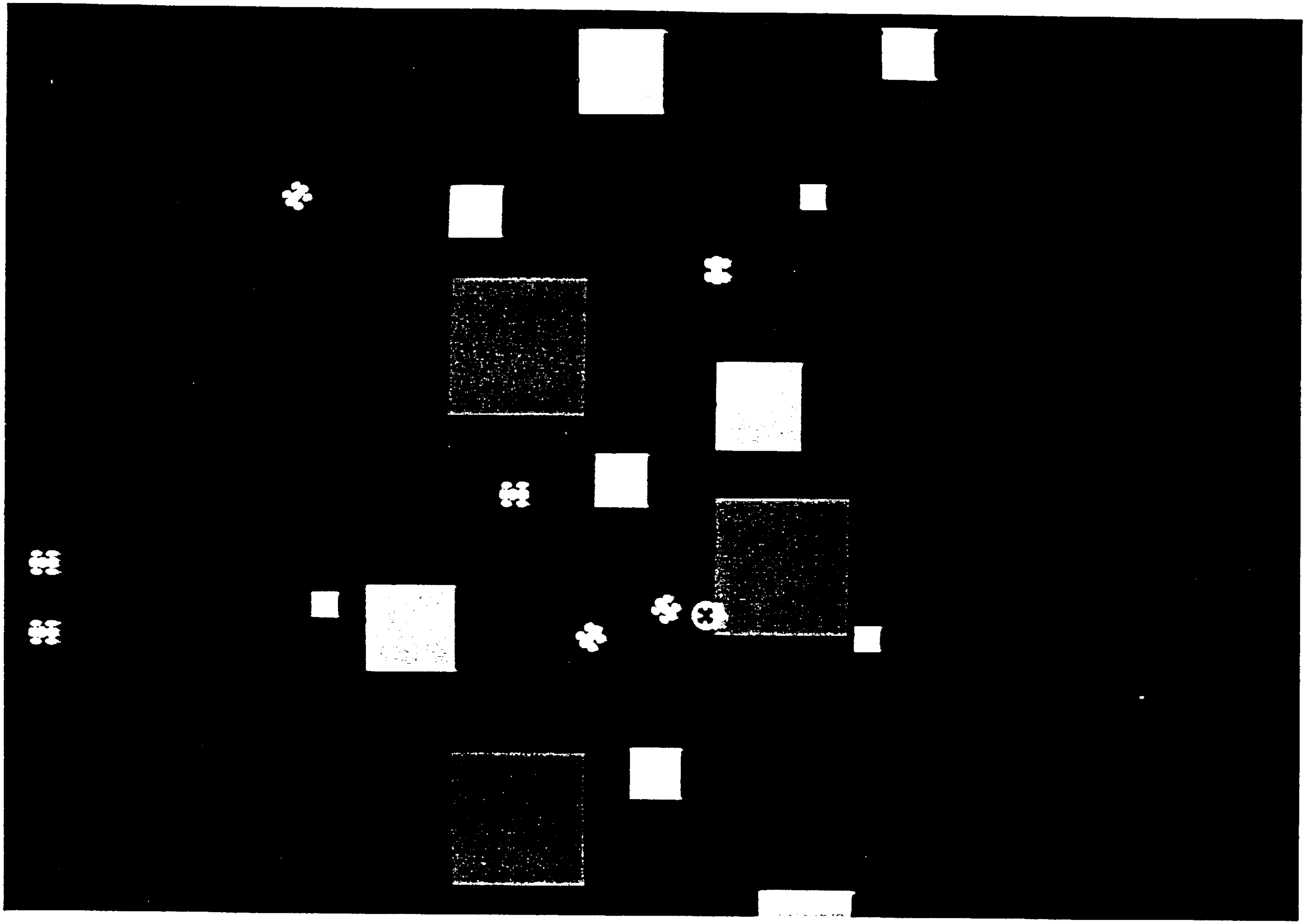
QUALITY-CONTROL INSPECTION
OASIS user looks at incoming parts, at own pace
executes voice commands to accept or reject.



TELE-OPERATION OF ROBOTS
OASIS user directs grappler by looking at desired location
executes voice commands to pick up and drop objects into box.



AIR-AIR OR GROUND-AIR FIRE CONTROL
OASIS user unconsciously identifies target
executes fire control by saying "Fire."



SIMULTANEOUS CONTROL OF SIX VEHICLES
OASIS user unconsciously identifies vehicles to be controlled
executes control commands using speech.

In the quality-control application (Figure 4-1), three conveyor belts move wrenches onto and off of the screen. Wrenches enter the scene at the left and, in the absence of user intervention, continue off the screen at the right. An X-shaped feedback cursor is presented continuously, indicating the user's focus of attention. Two input words are used to control the sorting process: "DOWN" and "UP". When the user utters a command, the feedback cursor blinks to indicate that the command was received. When a wrench is successfully designated, its image is flashed blue. If a wrench is selected with the "DOWN" command, it is transferred to the bottom conveyor belt and moves off the screen. Likewise, if a wrench is selected with the "UP" command, it is transferred to the top conveyor belt and moves off the screen.

In the teleoperation application (Figure 4-2), a collection of stationary objects scattered near the left edge of the screen must be picked up and placed in the hopper on the right. The feedback cursor appears as a disembodied robot grappler. Displayed continuously, it indicates the user's focus of attention. Two input words are used to control the grappler: "GRAB" and "DROP". When the grappler is over the object of interest, the user enters the "GRAB" command. If the designation is successful, the grappler closes and the object begins to move with the grappler. When the object has been positioned above the hopper, the user enters the "DROP" command. Upon receipt of the "DROP" command, the grappler immediately opens and releases the object which falls in a straight line into the hopper or disappears off the bottom edge of the display if not correctly positioned.

In the fire control application (Figure 4-3), target helicopters move at random across the screen. A circular cross-hair feedback cursor is presented continuously, indicating the user's focus of attention. A single input word is used to execute fire control: "FIRE". When the user utters the "FIRE" command, the feedback cursor blinks to indicate that the command was received. If the targeting was successful, there will be a slight delay and then the helicopter will flash and disappear from the screen.

In the vehicle control application (Figure 4-4), six terrain rovers must be steered through an obstacle course. Initially, all six rovers are stationary and positioned to the left of the screen. The obstacles, stationary squares and rovers, are distributed across the center of the screen from top to bottom. A circular cross-hair feedback cursor is presented continuously, indicating the user's focus of attention. Five input words are used to control the rovers: "LEFT", "RIGHT", "STOP", "GO", and "REVERSE". When the user utters a command, the feedback cursor blinks to indicate that the command was received. If a rover is successfully designated with the "LEFT" command, it changes its heading 45 degrees to the left, and so on. When a rover encounters an obstacle, it halts and can only be extracted using the "REVERSE" command.

Since all of the application demonstrations are programmed in PLEX, they can be readily reconfigured. A PLEX program can be edited in a matter of minutes. The form, color, and visibility of the feedback cursor can be changed by modifying a single line of text. Successful designation ranges, velocities of task objects, and choice of eye data filter can be changed as easily. With minimal programming, new features such as user control of cursor characteristics and eye data filters could be implemented. In addition to showcasing eye-voice applications, the existing demonstrations also testify to the flexibility and power of the development environment provided by the OASIS testbed.

5. SUMMARY AND CONCLUSIONS

This section lists and describes our various efforts towards establishing the next steps for OASIS. These efforts include:

- Summary of typical OASIS applications,
- Description of current OASIS-related research at Analytics,
- Three OASIS concept papers,
- Listing of public- and private-sector OASIS presentations, and
- Listing of OASIS magazine/newspaper articles.

Each of these efforts is described below.

5.1 SUMMARY OF TYPICAL OASIS APPLICATIONS

The potential application areas listed below are exemplary areas, by no means all-inclusive. They represent the thinking of the project staff and Analytics as to which applications seem most likely, given the research conducted during the two years of the OASIS project.

5.1.1 Government

- SINGLE-CREW COMBAT HELICOPTERS
  -- FIRE CONTROL
  -- FLIGHT CONTROL

- ADVANCED TANKS
  -- FIRE CONTROL
  -- TURRET CONTROL
  -- ONE-MAN TANK MANAGEMENT

- BATTLEFIELD ROBOTIC SYSTEMS
  -- UP TO SIX SIMULTANEOUS VEHICLES OR MISSILES
  -- MANEUVERING
  -- TARGET ACQUISITION AND DESIGNATION
  -- FIRE CONTROL

- TARGETING UNDER COMBAT CONDITIONS
  -- NEARLY INSTANTANEOUS RESPONSE
  -- OBSCURED VIEW
  -- MULTIPLE TARGETS

- INTELLIGENCE AND TACTICAL DATA FUSION
  -- GRAPHICAL COMMAND-CONTROL
  -- IMAGERY ANALYSIS

5.1.2 Commercial/Private Sector

- ROBOT CONTROL AND TRAINING

- CONTROL OF MULTIPLE ROBOTS IN FACTORY

- PARTS INSPECTION

- MICROSURGERY

- ROBOT IN RADIOACTIVE ENVIRONMENTS
  -- WASTE DISPOSAL
  -- REACTORS

- SYSTEMS FOR THE HANDICAPPED
  -- ELECTRONIC CONTROLS
  -- READING AND WRITING AIDS
  -- WHEELCHAIRS

- AIRCRAFT CONTROL AND SAFETY
  -- FLIGHT DATA RECORDING
  -- TRAINING
  -- NAP-OF-EARTH MANEUVERING

5.2 OASIS-RELATED RESEARCH AT ANALYTICS

There are two just-completed SBIR Phase I projects at Analytics with OASIS-oriented themes. The first (Arnold et al., 1986) studies the feasibility of recording pilot eye movement data in space or commercial flight. The second (Harrington et al., 1986) deals with computer control via brain wave plus ocular data. For each project, Phase II proposals have been written and submitted. Brief summaries of each project are given below.

5.2.1 Recording Pilot Eye Movement Information on a Digital Flight Data Recorder. Phase I SBIR Technical Report delivered September 1986 to Dr. R. Harris (COTR), NASA Langley Research Center.

Summary

The purpose of this study was to determine the feasibility of an innovative concept to collect data in the cockpit on pilot eye movement using an oculometer and record that data on the Digital Flight Data Recorder (DFDR) for subsequent analysis. The feasibility study focused on three areas: (1) an investigation of current state-of-the-art oculometer hardware; (2) an investigation of the technical issues regarding data recorders (their signal characteristics, interface requirements, and preprocessing requirements); and (3) a study of the utility of pilot eye data for accident investigation and general understanding of cockpit human factors issues. The research methods used included literature review, analysis, personal communication, and direct meetings with the National Transportation Safety Board and Federal Aeronautics Administration.

The completion of the Phase I research resulted in the following conclusions:

- The concept is feasible, and the development of a prototype system is warranted.
- Oculometer components exist that appear suitable, with modification, to the cockpit environment.
- The DFDR has the capacity and flexibility to record processed visual data.
- The development of processing and control algorithms is feasible.
- The information recorded on the DFDR would have extensive utility in accident investigation and other areas of interest.

The successful development of this concept will benefit NASA and the commercial aviation industry in several ways. The successful capture of pilot visual data will have application to both space and atmospheric flight in regard to increased safety by allowing a better understanding of the human issues surrounding in-flight accidents. Additionally, the data acquired through this concept will support research on pilot scan patterns, workload, stress effects, and piloting techniques. This will result in improved cockpit performance, flight safety, and better understanding of a complex man-machine interface.

5.2.2 Magnetoencephalography for Real-Time Computer Control. Phase I SBIR Technical Report delivered June 1986 to Mr. James Villareal (COTR), NASA Johnson Space Center.

Summary

The purpose of this study was to determine the feasibility of an innovative concept that uses an operator's brain waves as a control mechanism for computer systems. The feasibility assessment was based on both the latest advances in brain wave sensing technology as well as the unique control requirements and characteristics of intelligent computer systems. The results of the feasibility assessment indicate that, at the present time, brain wave recording technology is not adequate for data transmission; however, the field of biomagnetism is advancing at a rapid rate. Sophisticated hardware is currently under production, and dramatic enhancements are expected in software developments as more extensive multiple-sensing systems are introduced.

The potential application of this research is the development of a system for monitoring operator states. A system that records and correlates human ocular and brain wave activity has utility in any environment where the operator is required to interpret information, analyze information, and make decisions. An intelligent system that would function in a cooperative role with the operator could reduce the operator's workload and improve job performance.

Further development of this concept requires a precise understanding of the linkage of three components: operator eye movement information, brain wave activity, and task structure. Analytics has developed testable hypotheses that address the issues of eye/brain/task linkage in the "computer control" context. Analytics has developed design goals, functional requirements, system architecture, and equipment specifications for an eye/brain/task (EBT) testbed. The development of the EBT testbed and the investigation of the experimental hypotheses are the core of the Phase II proposal.

5.3 OASIS CONCEPT PAPERS

This subsection contains summaries of three concept papers written by Analytics. Each paper takes a specific OASIS application, explores the ideas, and then proposes an approach to further investigate the ideas. The three concepts are:

1. Determination of Pilot Intent
2. Monitoring Pilot Consciousness
3. Aids for the Handicapped

5.3.1 Automatic Determination of Pilot Intent Using OASIS. Concept paper submitted to Lt. Col. John R. Retelle, Program Manager, Tactical Technology Office, USAF.

Summary

Extensive efforts have been devoted to automation of fighter cockpit functions in order to achieve a highly capable single-pilot fighter which does not impose excessive demands on pilot performance. A major obstacle to achievement of this goal has been the bottleneck in communications between pilot and system. While each new automated component performs important information-processing functions, it also tends to generate new tasks for the pilot in monitoring and controlling that component.

Several converging factors create both the critical need and the immediate opportunity for a major advance in pilot-cockpit interface:

- Present pilot automation systems do not measure pilot intent; they rely on manual pilot inputs which may or may not be available in a timely fashion.

- A good copilot does measure and respond to pilot intent.
- Pilot workload is increasing as more sophisticated systems come into play.
- Even if the presence of an additional crewmember were possible to absorb the increased weapons management and related workloads, adding such a crewmember would cost 500-1000 lbs. of empty weight and up to 10,000 lbs. of gross weight per aircraft; conversely, removing a crewstation would save these very significant amounts.
- Modern "soft" displays offer new opportunities to provide optimal responses to pilot intent, once measured, in terms of efficient displays of vital data and therefore more rapid human reactions.

In principle, then, an automated copilot able to predict intent will have large potential payoffs in platform and human performance. The OASIS system, coupled with other current intelligent system technologies, may offer the required capability.

It is desirable for cockpit automation to be manifested in the form of an automated copilot, an intelligent system that understands the pilot's goals and takes direction from his actions. A good human copilot can casually observe the pilot and take action cues from subtle glances and gestures. For example, the copilot may notice that the pilot is looking at the altimeter on an approach to landing and may guess that the pilot would like the landing gear to be lowered; the copilot may grasp the landing gear lever and look for a confirming nod from the pilot before proceeding with the action. An automated copilot should be equally unobtrusive, but with additional capabilities to control the format and content of digital cockpit displays, presenting the pilot with just the information that he needs at each moment.

The chief problem in developing an automated copilot has been the difficulty of enabling the system to "see" enough of the pilot's behavior to act intelligently and with initiative; the conventional cockpit can sense pilot actions only in the form of inputs that are made using manual control devices. It is now possible to use advanced technologies to design an automated copilot which can truly see and understand the behavior of the human pilot.

Specifically, under contract with NASA, Analytics has demonstrated the feasibility of using automated speech recognition and eye movement monitoring in conjunction with signal processing to determine the foci of the pilot's visual attention and to correlate verbal communications relating to the visual foci.

5.3.2 Pilot Consciousness Monitoring Using OASIS. Concept paper submitted to Lt. Col. John R. Retelle, Program Manager, Tactical Technology Office, USAF.

Summary

Fighter aircraft are incorporating a broad range of automation and man-machine interface technologies to expand and enhance the capabilities of the total human-machine weapons system; this is indeed the goal of the Pilot's Associate Program. The human role is changing from that of a real-time controller to one of a mission manager, with increasing demands being imposed on human cognitive and psychomotor capabilities. At the same time, the flight performance capabilities of advanced fighter aircraft have become sufficient to threaten the physiological survival of the human pilot under various manually or automatically directed maneuvers; for example, high sustained G levels and high G onset rates can easily render the pilot unconscious. Pilot incapacitation from G-induced loss of consciousness (G-LOC) or other causes (e.g., disorientation, decompression, seizure, etc.) has been responsible for the loss of many aircraft and pilots in the past (Rayman, 1973) and could potentially continue to pose equally catastrophic conditions even with the pilot assuming more of a mission manager role. It is important, therefore, to devise a way to use automation and interface technologies to ensure the integrity of the pilot-vehicle system under all plausible operational conditions.

It is currently possible to construct autopilot systems which can effectively assume aircraft control and achieve a stable flight path in the case of temporary pilot incapacitation. While it is difficult to find reliable sensors for assessing pilot consciousness, it is both feasible and valuable to implement such an autopilot recovery system based solely on a minimum-altitude trigger; in fact, such a system has been demonstrated for the Air Force

AFTI/F-16 (Howard and Johnston, unpublished). Although this system is satisfactory for many conditions, it is not applicable where low-altitude flying and maneuvering under manual control are required; since many warfare scenarios call for low-level fighter penetration, this deficiency is critical. Further, a recovery system based on a minimum-altitude trigger might be inadequate for conditions in which the pilot is flying far above the minimum level so that autorecovery might be delayed too long.

In order for an automatic recovery system to operate at low altitudes, it would have to detect pilot incapacitation directly and reliably and be able to take control quickly. It should additionally be unobtrusive, fail-safe, and make maximum use of existing cockpit equipment (i.e., imposing a minimal demand for new specialized equipment in the cockpit). Finally, it must provide for immediate pilot override.

It is possible to use eye movement monitoring to detect pilot incapacitation caused by G-LOC. It may additionally be possible to detect pilot incapacitation deriving from other causes using the technique described in Glenn (1986a) for determination of pilot intention with OASIS. Studies which have imposed high G conditions on subjects in centrifuge experiments have repeatedly demonstrated a characteristic eye fixation behavior which reliably precedes G-LOC by a few seconds (Beckman et al., 1961; Coburn et al., 1963). Such a system could be based on the use of the OASIS concept for eye-voice control of interface functions which is already envisioned for aircraft cockpit implementation (Glenn et al., 1984; Glenn, 1986a). The recovery system could also incorporate elements of state-of-the-art systems for terrain following/ terrain avoidance (TF/TA) in order to ensure a rapid and safe transition to a default flight path (e.g., level flight at preset altitude, course, and speed) while awaiting pilot recovery.

5.3.3 <u>Application of Eye Position and Voice-Driven Computer Interface to Aid the Physically Disabled. Phase I SBIR proposal submitted to the National Science Foundation and to the Department of Education</u>.

Summary

A large portion of the physically handicapped population (e.g., quadriplegics) are intelligent, motivated, and trainable. These attributes can allow them to become productive members of society, but physical disabilities prevent their utilization of modern computer-based tools, training systems, and vocations.

A variety of man-machine interface concepts have been developed in recent years in an attempt to: (1) increase the flow of relevant information between the system and operator, and (2) alleviate the need for complex, programmer-oriented inputs through the use of user-friendly workstations. Unfortunately, for the severely physically disabled, these interface concepts usually rely on the keyboard as the human-to-computer input device. Keyboard technology, no matter how sophisticated and user-friendly, normally requires the use of the operator's hands; this requirement excludes the severely physically handicapped.

Automated voice recognition systems provide an innovative interface device which does not require manual intervention. However, voice technology is not sufficient in itself to solve the problem due to three obstacles:

1. The technology has many technical limitations such as vocabulary limitations, extensive training requirements, etc.

2. Voice input is not conducive to the input of continuous, locational information.

3. Many severely disabled people have impaired voice control.

A requirement then exists for a means to increase the power and range of the voice actuation device, overcoming vocabulary limitations yet providing the control system operator with hands-off manipulation. We believe this requirement can be met through the use of a non-intrusive eye movement recording instrument (e.g., an infrared oculometer) in conjunction with a voice-actuated display system. We propose to adapt the results obtained in the OASIS effort to build a generic eye movement and voice-controlled interface for use by severely physically disabled people.

5.4 LISTING OF OASIS PRESENTATIONS

This section lists the OASIS presentations made to various groups in the U.S. government, especially the Department of Defense and NASA. Presentations were made primarily by personal visit to the OASIS Lab/Testbed or by showing an OASIS demonstration videotape.

Presentations were also made to a number of private-sector organizations not listed here. Though no commitments have yet been made, most of the presentations were received with considerable interest. Impediments to commitments were reduced budgets for basic research at almost all government agencies and lack of available funding vehicles.

Client	Agency
U.S. ARMY	
Dr. Charles Church	DCSRDA, Pentagon, Washington, DC
Eugene DelCoco	USA-ARDEC, Dover, NJ
Kennard Raisner	USA-ARDEC, Dover, NJ
Dr. Daniel S. Berliner	U.S. Army Medical R&D Command, Ft. Detrick, MD
C. Tsowbanos	AVSCOM, St. Louis, MO
J. Lane	U.S. Army TACOM, Warren, MI
Dr. R. Lighty	U.S. Army ETL, Ft. Belvoir, VA
Lt. Col. G. Downs	AATD, Ft. Eustis, VA
Capt. L. Campbell-Wade	APTD, Ft. Eustis, VA
J. Respass	AVRADCOM, Ft. Monmouth, NJ
H. Cohen	AMSAA, Aberdeen Proving Ground, MD
Lt. Col. J. Alexander	U.S. Army Technical Integration Office, Vienna, VA
Mr. Clarence Fry	U.S. Army Human Engineering Lab, Aberdeen Proving Ground, MD
Dr. D. Hislop	LABCOM/HDL, Adelphi, MD

Client	Agency
OTHER DoD	
Dr. G. Calhoun	USAF Aerospace Medical Research Lab, Wright Patterson AFB, OH
Cdr. W. Moroney	Naval Air Development Center, PA
Lt. Col. J. Retelle	DARPA, Tactical Technology Office, Rosslyn, VA
Dr. R. Engle	OSD/C^3I, Pentagon, Washington, DC
NASA	
Dr. R. Harris	Langley, VA
Mr. R. Courtney	HQ/Code S, Washington, DC
Dr. M. Montemerlo	HQ/Code R, Washington, DC
Dr. S. Ellis	Ames Research Center, CA
Dr. L. Allen	JPL, CA
Ms. J. Brown	Johnson Space Center, TX
OTHER GOVERNMENT	
Mr. Jack Ryan	FAA, Air Traffic Operations Center, Washington, DC
Dr. W. Shulthies	CIA, Washington, DC

5.5 OASIS NEWSPAPER/MAGAZINE ARTICLES

The final exhibit in this section comprises six articles written on OASIS from the *Wall Street Journal, Science '86, Philadelphia Magazine, Financial Times, Computerworld*, and *P.M. Magazin* (Germany).

THE WALL STREET JOURNAL.

EASTERN EDITION TUESDAY, APRIL 1, 1986 PRINCETON, NEW JERSEY

What's New: Eye Commands, Sprays, Bombs

AN INDIVIDUAL'S EYES may indeed be the windows to the soul of a new machine.

Analytics Inc. in Willow Grove, Pa., has developed a computer that can carry out voice orders to act on what a person is looking at. The company's Ocular Attention-Sensing Interface System, or Oasis, may eventually be used to control aircraft on radar screens or in systems for the disabled or in microsurgery.

The National Aeronautics and Space Administration has funded much of the development of Oasis, in hope that it will lead to a more direct connection between a pilot and the computers that control flight. "A man flying an airplane these days is really flying a computer," says Randell Harris, a NASA physiologist. "It would be fairly natural for someone to look at something, then give a computer a command to do something to it."

This is one of an occasional series of reports on technology that has reached the early stages of application.

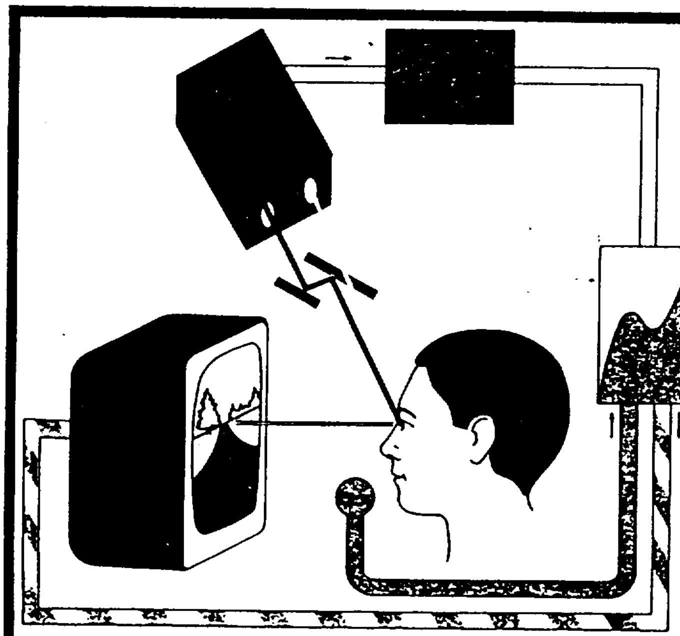
Natural, but not that easy. A subject testing Oasis sits at a color monitor while an infrared beam is trained on his right eyeball. Once calibrated, the computer tracks the eyeball, using the beam's reflection. As a series of enemy helicopters crosses the screen, the subject focuses on one aircraft and gives the command, "Fire." If eye and machine are in sync, the target is hit.

A major difficulty is the eye's tendency to wander, Even in something as simple as approaching a traffic light, the eye is constantly and rapidly moving among three signals—not just one. Although the accuracy of Oasis has varied with the user, Analytics expects to have within two years a production model that can respond to a one-inch shift in focus at 30 feet.

SCIENCE 86

APRIL VOLUME 7, NO. 3

CURRENTS



Effortless computing: the eyes have it

WILLOW GROVE, PENN.—Computing has never been so hands-off. An electronics firm, Analytics, Inc., has developed a computer system that works primarily under guidance from the operator's eyes. Other research teams have developed systems with a helmet-mounted apparatus that relies on head movements to guide the blinking blob of light, or cursor, that marks your place on a computer screen. But the Analytics system needs no headgear. As shown at left, a user watches objects—displayed in this case on a video screen. A tightly focused infrared beam (red) shines into one eye and bounces back (yellow) to a camera. Sixty times each second, the camera records the location of the pupil relative to the cornea—two points that show where the operator is looking. Extraneous eye movements, such as blinks and twitches, are filtered out (orange box). By focusing on a target and then giving a simple verbal command (green) into a microphone, a person can use the system to control a robot, track enemy aircraft, or sort parts on an assembly line. Analytics president Stephen Leibholz says the company hopes to market the system to industry and the military by 1988.

STEVE BASHARA STANSBURY RONSAVILLE WOOD INC

Philadelphia
MAGAZINE

May 1986 Vol. 77 No. 5

Tech

A COMPUTER THAT'S MORE THAN MEETS THE EYE: A Willow Grove electronics firm has developed a new kind of computer system that is taking the joy stick out of computing—a program that operates primarily under guidance from the operator's *eyes*. **Analytics, Inc.**'s new system integrates two technologies, one that measures eye movement and one that recognizes voices, so that a user need only direct his eyes to one point—say on a computer screen—and speak a command into a microphone (like "grab," "fire," or "drop"), and the command will be carried out. The system works with a tightly focused, low-level infrared beam that shines into one eye and, 60 times a second, determines exactly where the eye is looking (by analyzing the disparity between the location of the pupil and the cornea). Project manager **Dr. Allen Zaklad** says the system, called **Oasis,** may be used to control robots, track enemy aircraft, sort parts on an assembly line, or direct laser beams in surgery. —*Rachael Migler*

FINANCIAL TIMES

EUROPE'S BUSINESS NEWSPAPER

No 29,907 Friday April 18 1986

EYE MOVEMENT is being used by Analytics of Willow Grove, Pennsylvania, in an experimental computer system that allows the user to look at an item on the screen and speak a command that will act on that item. For example, he might look at a segment of graphics, utter the word "blue," and the segment turns blue.

An infrared beam is trained on one eyeball. Once calibrated, the computer can monitor eyeball movement, and therefore direction of glance, by measuring the reflected movement of the beam. The company expects to have a production model in two years.

THE NEWSWEEKLY FOR THE COMPUTER COMMUNITY

COMPUTERWORLD

$2/COPY; $44/YEAR NOVEMBER 3, 1986 VOL. XX, NO. 44

Researchers focus on promise of eye-gaze technology

By ALAN ALPER

In the 1977 science-fiction novel *Firefox*, the Soviet Union has developed a fighter plane that can fly six times the speed of sound — undetected by radar — and that has an integrated weapons system that can be operated by thought waves.

Sounds farfetched? Not really.

Today, in a number of small university and corporate labs, technology is under development that would enable a human being to control a computer, navigate an airplane or maneuver a robotic system by eye movement. While the focus, for the most part, has centered on how to make life easier for the physically handicapped or how to actually develop an aircraft that could be operated by gazing at flight instruments, work is being done to adapt this technology to a variety of other applications, including data processing.

While the keyboard, mouse and joystick will continue to be the primary ways to control a computer, many contend the human eye gaze will one day become an invaluable adjunct to these methods. Some believe research in eye-gaze technology could shed light on the ultimate control mechanism — the brain.

"We believe that eye gaze is just another avenue to increase the man/machine bandwidth," says Gary Kilikany, vice-president at Sentinent Systems Technology, a Pittsburgh start-up that recently unveiled a system that enables people who cannot communicate to do so by controlling a computer with their eyes.

"In some situations, there's nothing more natural than looking at something like a video screen to select menus or to push buttons," Kilikany says. "It's a very natural alternative to a man/machine input mechanism."

IBM scientists at the Thomas J. Watson Research Center in Yorktown Heights, N.Y., spent a good deal of time in 1981 researching eye gaze as a means of controlling a computer. IBM, which was recently awarded a patent for an eye-tracking mechanism, was attempting to develop a high-resolution display that could be controlled by eye movement.

"We never finished the project," recalls Jim Levine, a scientist on IBM's research staff, noting that the pre-IBM Personal Computer program used a Series/1 minicomputer, which would have made a commercial product extremely expensive to purchase. "We did build an eye tracker, however, that was accurate enough to control a computer."

While IBM dropped the project soon thereafter, the fruits of that labor are being enjoyed by researchers at the University of Virginia in Richmond, Levine says. "They are working on an eye-tracking system for the handicapped that they hope to soon build into a product," he remarks. "We've loaned them some PCs and are doing some consulting on the project."

Labs like the one at the University of Virginia and another in the Trace Research & Development Center on Communication, Control & Computer Access for Handicapped Individuals at the University of Wisconsin at Madison's Waisman Center continue to push the technology to its limit.

Perhaps the most ambitious research, however, is being done by a small defense contractor nestled in the hills of northeast Pennsylvania.

There, a 19-year-old privately held firm, Analytics, Inc. in Willow Grove, has spent the last 18 months developing eye-gaze technology, used in concert with existing voice-recognition systems, to control computers, robots and vehicles. The development work is being financed primarily by the National Aeronautics and Space Administration under the Small Business Innovative Research Program.

Visual attention, vocal intention

Called the ocular attention-sensing interface system (OASIS), the device measures an operator's visual attention and vocal intention.

"We are looking at the foveae — where a person sees — to get attention and are coupling that with speech recognition to get intention," notes Analytics' President Steve Leibholz. "It's the closest thing to automatic or unconscious control."

OASIS uses a technique in which light is projected into one eye, a portion of which is reflected by the cornea to create a virtual image that responds to changes in the relative position of the eyeball. Using a high-speed analog signal processor, OASIS is said to monitor movements of the eye of less than 1 degree of arc.

Also used is a Texas Instruments, Inc. voice-recognition system in which a series of single-word commands are stored.

Eye movement and voice data are sent to five algorithmic modules that analyze the eye movements and voice patterns, among other things, and translate them into system commands.

In current studies, a subject sits before a color monitor in a stationary position, and an oculometer is fixe on his pupil and cornea to follow th eye's movement. The subject's voic commands are stored in the speec recognizer.

The subject is then asked to follo the movement of objects on the mon tor and to invoke the command "fire when the cursor becomes synchr nized with the target. If the cursor i in synch when the command is give the target is destroyed.

While NASA is most interested i the technology for its manned spac flights and stations, Analytics co tends the OASIS's applications ar endless. The company has alread come up with 25 applications, incluc ing aiding the handicapped, air tra fic control, robotics and compute system management.

Under the computer heading, the firm lists data retrieval, computer-aided-design and manfacturing, photographic interpretation, signal processing, supercomputer process management and computer vision.

While viewed as an adjunct to other input and control methods, Leibholz foresees OASIS being used as an extremely fast and accurate way of searching through dense data bases for desired data. Some view it as an Evelyn Woods approach to computing.

"In our concept, an operator can pick one item out of 100 items that flash by on screen 100 times as fast as if doing it manually," Leibholz says. "It's a more sophisticated way of visual recognition."

It is the voice portion of OASIS that may prove to be difficult to adapt to commerical settings, he adds, noting the limited number of words such devices can recognize as well as the ambient noise as factors that cause inaccuracies.

Analytics must overcome other technological hurdles as well. Natural occurrences, such as blinking and eye drift, cause OASIS some eye-tracking problems. Also, in its current stage of development, corrective lenses and sudden head movements throw OASIS off.

"People have a tendency to move their eyes suddenly — it's a means of defense," IBM's Levine says. "Something like that is always there and is hard to overcome."

So far, under the NASA Small Business Innovative Research program, Analytics has received $550,000 to prove the feasibility of its concept, much of which has been spent acquiring equipment.

The firm is currently seeking additional funding to develop two engineering prototypes in 1987 that will be firmware and software driven and will use multiprocessor technology. Leibholz says the firm is looking for investors from the private sector but is also considering breaking out the OASIS project as a separate company via a public offering.

Meanwhile, Analytics hopes knowledge gleaned through development work on OASIS on how the brain analyzes information received from the eye will form the foundation for study of actual thought-controlled systems. Leibholz believes that by using a noncontact magnetoelectroencephlograph to measure brain activity, inferences can be made about what a person is thinking.

"There's no reason why you can't get at what a person is thinking," Leibholz says, "I'm not prepared to go into detail, but there is potential feasibility using a magneto-electroencephlograph to achieve knowledge of attention of focus or some measure of control."

While development work at Analytics continues, Sentinent is already marketing a device that uses eye-gaze technology to enable physically impaired people to communicate.

The firm recently began shipping a cost-reduced version of its eye-tracking device — called Eyetyper — that is priced at $3,000 and can be connected to microcomputers through an RS-232 port. The 3-year-old company was founded by former Carnegie-Mellon University engineering students in Pittsburgh who were involved in a volunteer project to enable children with cerebral palsy communicate with their eyes.

So far, Eyetyper is mainly used in intensive care units, rehabilitation hospitals and special education school systems. Kilkany, however, sees many potential applications in the not-too-distant future for the technology.

"Right now, the system is used by people who can't speak or move their hands or legs easily," Kilkany says. "It's not sold to mainstream America — at least not yet. We do want to go in that direction.

"I can see it used in computing as a device that enables a user to select menus," he continues. "Or, in factories where a worker's hands are busy, and he needs to register defective parts and can do so by looking into an LED." ■

PHOTO COURTESY OF SARAH SMITH

Analytics' Steve Leiboltz

P.M. Magazin
19 September 1986

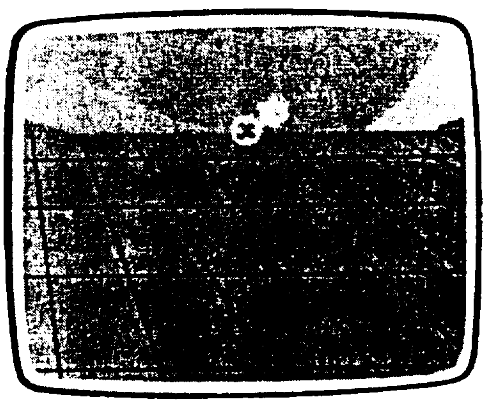
Infrarot-Technik

Wie Augen eine Maschine steuern können

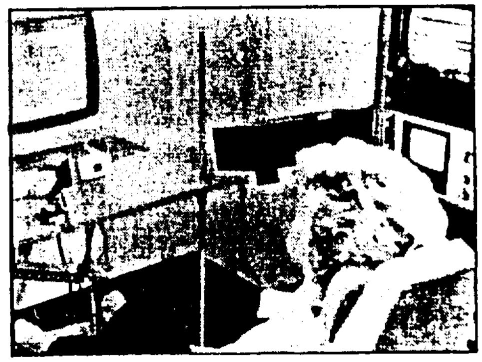
Klingt wie Science-fiction: Man hockt vor einem Bildschirm, blickt auf ein darauf abgebildetes Haus, sagt »Hinfahren«, und schon ist das Fahrzeug auf dem Weg dahin. Ein Gerät, das sich zu einem solchen System ausbauen läßt, hat jetzt eine Computerfirma in den USA vorgestellt. Obwohl es zunächst rein militärischen Zwecken vorbehalten ist, zeichnet sich bereits jetzt ab, daß es später Behinderten und Ärzten wertvolle Hilfe leisten wird. Nachdem es Computer, die auf Sprachanweisung gehorchen, schon eine Weile gibt, ist die Beeinflussung durch die Augen das Neue an diesem Gerät.

Das geht über einen auf das rechte Auge gerichteten und von ihm zurückgeworfenen Infrarotstrahl. So läßt sich jede Bewegung des Augapfels registrieren und sich anschließend auf ein Koordinatensystem übertragen. Sobald jetzt ein bestimmter Bildgegenstand fixiert wird, schwenken Fahrzeug, Waffe oder Skalpell in die vom Auge vorgegebene Richtung und führen sodann den ausgesprochenen Befehl aus.

Soweit die Wunschvorstellung. Kopfzerbrechen bereitet unterdessen noch, daß, um hinreichend präzise arbeiten zu können, winzig kleine Richtungswechsel des Auges ermittelt werden müssen. In zwei Jahren hoffen die Hersteller mit einem Modell antreten zu können, das eine Augenbewegung erfaßt, die auf einer zehn Meter entfernten Wand fünf Zentimeter betragen würde. ★



Anfahren des Ziels: Die beiden Markierungen verraten dem Navigator seine Entfernung vom Ziel. Sobald er den hellen Punkt fixiert, bewegt er sich mit dem dunklen darauf zu.



Gelenkt wird mit den Augäpfeln: Ein feiner Infrarotstrahl tastet das rechte Auge ab und überträgt seine Bewegung auf den Bildschirm.

5.6 CONCLUSIONS

Phase II research efforts proved the feasibility of OASIS as an innovative interface option for human-system interaction. A highly flexible and experimentally powerful OASIS testbed was built, including a graphics editor and a data collection and analysis system. OASIS algorithms were highly parameterized and could be dynamically modified in real time for development purposes. Four demonstration graphics tasks were produced which simulated quality-control inspection, teleoperation of robots, air-air or ground-air fire control, and simultaneous control of six vehicles. The tasks were performed with voice and eye control only. The testbed laboratory was also used to conduct an exploratory multi-factor experiment which investigated alternative OASIS eye-processing algorithms and visual feedback modes for a task involving designation of stationary objects. In addition, the initial experiment compared OASIS performance with that of a mouse.

The experimental results showed that OASIS compared favorably with the conventional rapid-pointing mouse interface for both acquisition and stabilization times. The OASIS EPT eye filter, combined with non-continuous visual feedback, resulted in the best task performance. However, these experimental results cannot be generalized to all tasks and applications. The optimum set of OASIS parameters depends strongly on the characteristics and demands of the final application task. This became apparent during the development of the demonstration samples. The parameter set which worked best for designating stationary objects differed from that for designating moving objects.

Ideally, further experimentation and OASIS optimization will proceed after the selection of a specific application. Using the highly flexible graphics and icon editor built for the OASIS testbed, a laboratory simulation can be rapidly developed. An initial set of experiments can then be designed which would investigate key components of the OASIS interface relevant to the targeted system and targeted user population. Thus an OASIS experimental program would be fully integrated with the development of the new system.

ACKNOWLEDGMENTS

The work reported here was performed under NASA's Small Business Innovative Research (SBIR) program (Contract Number NAS7-932). Monitoring of the contract was performed by the NASA Resident Office at the Jet Propulsion Laboratory (JPL). We gratefully acknowledge the assistance we received from Mr. Peter Tackney, Contracting Officer, and Mr. Marvin Perlman, Contracting Officer's Technical Representative, at JPL.

Many people at NASA centers offered their assistance to the OASIS project, especially Dr. Randall Harris (Langley), Dr. Mel Montemerlo (HQ), and Dr. Steve Ellis (Ames). Dr. Felix Barker and his associates at the Pennsylvania College of Optometry provided special clinical expertise. We also want to thank the two major equipment vendors -- Applied Science Laboratories (oculometer) and Masscomp (graphics and system integration) -- for timely help in troubleshooting our system.

Finally, we wish to acknowledge the many people at Analytics who helped us reach a substantial degree of success in building OASIS. Bill Weiland, Lorna Ross, Dan Leibholz, and Dan Weiss played important roles in designing and building the software. Steve Rodgers and Ed Kapnic helped us to bring OASIS into the public eye. Steve Leibholz, the President of Analytics, contributed both technical ideas and inspiration to the project.

REFERENCES

Arnold, W., Deimler, J., and LaGrossa, C. Recording Pilot Eye Movement on a Digital Flight Recorder. Analytics Technical Report 2042, Analytics, Inc., Willow Grove, PA, 1986.

Beckman, E., Duane, T., and Coburn, K. Limitation of Ocular Motility and Pupillary Dilatation in Humans Due to Positive Acceleration. NADC Technical Report NADC-MA-6140. Johnsville, PA: Aviation Medical Acceleration Laboratory, Naval Air Development Center, 1961.

Coburn, K., Beckman, E., and Duane, T. "Limitation of Ocular Motility under Acceleration: An Objective Endpoint for Use with Human Centrifuge Subjects." *Aerospace Medicine*, September 1963, pp. 824-825.

Eizenman, M., Frecker, R., and Hallett, P. "Precise Non-Contacting Measurement of Eye Movements Using the Corneal Reflex." *Vision Research*, 24(2), 1984, pp. 167-174.

Glenn, F.A. Automatic Determination of Pilot Intent Using OASIS. Analytics White Paper, Analytics, Inc., Willow Grove, PA, 1986a.

Glenn, F.A. Pilot Consciousness Monitoring Using OASIS. Analytics White Paper, Analytics, Inc., Willow Grove, PA, 1986b.

Glenn, F.A., Harrington, N., Iavecchia, H.P., and Stokes, J.M. An Oculometer and Automated Speech Interface System. Analytics Technical Report 1920, Analytics, Inc., Willow Grove, PA, 1984.

Glenn, F.A., Iavecchia, H.P., Ross, L.V., Stokes, J.M., Weiland, W.J., Weiss, D., and Zaklad, A.L. "Eye-Voice Controlled Interface." *Proceedings of the 1986 Human Factors Convention*, October 1986, pp. 322-326.

Harrington, N., Deimler, J., Stokes, J.M., Kaufman, L., and Williamson, S.J. Magnetoencephalography for Real-Time Computer Control. Analytics Technical Report 2035, Analytics, Inc., Willow Grove, PA, 1986.

Howard, J. and Johnston, A. AFTI/F-16 Gravity-Induced Loss of Consciousness and Spatial Disorientation Auto-Recovery System. Unpublished manuscript. Edwards AFB, CA: Air Force Flight Test Center.

Rayman, R. "In-Flight Loss of Consciousness." *Aerospace Medicine*, June 1973, pp. 679-681.

REFERENCES (Continued)

Stokes, J.M. VRAS System Guide. Analytics Technical Report 1400.19B, Analytics, Inc., Willow Grove, PA, 1982.

Wixon, S.E. "Four-Dimensional Processing Tools for Cardiovascular Data." *IEEE Computer Graphics and Applications*, August 1983.

Young, L. and Sheena, D. "Methods and Designs -- Survey of Eye Movement Recording Methods." *Behavior Research Methods and Instrumentation*, 1975, pp. 397-429.

Zaklad, A. Application of Eye Position and Voice-Driven Computer Interface to Aid the Physically Disabled. Proposal submitted to the U.S. Department of Education Application Control Center. Analytics Proposal No. 86-20045, Analytics, Inc., Willow Grove, PA, 31 March 1986.